随机微分方程及其应用概要

龚光鲁

清华大学出版社

随机微分方程及其应用概要

龚光鲁

清 华 大 学 出 版 社
北 京

内容简介

本书是为应用领域的读者撰写的关于随机微分方程的入门教科书.书中对于理论性概念的定义与命题的推导并不探求数学的严密性,而是通过剖析原始想法来叙述其含义及其可能的发展,使读者尽快地了解并掌握随机微分方程的思想要领,同时也为进一步学习、提高的读者提供了一个直观的平台.书中的内容安排对读者的知识准备要求较低,只需要具有初等概率论知识,而不要求具备测度论的知识.

本书适用于金融经济、系统工程、物理科学、系统生物学等领域中的教师、科研人员、管理人员、研究生、大学生作为教材或参考书,也可供有关人员自学之用.

图书在版编目(CIP)数据

随机微分方程及其应用概要/龚光鲁.—北京:清华大学出版社,2008.2

ISBN 978-7-302-16776-1

Ⅰ.随… Ⅱ.龚… Ⅲ.随机微分方程 Ⅳ.O211.63

中国版本图书馆 CIP 数据核字(2008)第 004130 号

责任编辑:刘 颖
责任校对:赵丽敏
责任印制:

出版发行:清华大学出版社
http://www.tup.com.cn
c-service@tup.tsinghua.edu.cn
社 总 机:010-62770175
投稿咨询:010-62772015

地 址:北京清华大学学研大厦 A 座
邮 编:100084
邮购热线:010-62786544
客户服务:010-62776969

印 刷 者:
装 订 者:
经 销:全国新华书店
开 本:185×230 **印 张:**15 **字 数:**307 千字
版 次:2008 年 2 月第 1 版 **印 次:**2008 年 2 月第1 次印刷
印 数:1~ 000
定 价: .00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换.联系电话:010-62770177 转 3103 产品编号:

前　　言

本书是为广大应用领域的同仁撰写的关于随机微分方程的引导性读物,适用于诸如金融经济、系统工程、物理科学、系统生物学等领域中想了解随机微分方程,或者需要用随机微分方程建模的教师、科研人员、管理人员、研究生、大学生作为教材或参考书,也可供有关人员用作自学的材料.

本书的主要内容是随机微分方程及其应用.在本书中,我们对于理论性概念的定义与命题的推导,并不探求数学的严密性,而是通过剖析原始想法叙述其含义及其可能的发展,让使用者尽快地了解并掌握随机微分方程的思想要领.同时也为想要进一步学习提高的读者提供了一个直观的平台.

使用本书并不需要测度论的知识,只需要有初等概率论的准备.例如掌握文献[QY]中的概念就可以完全通读本书.

随机微分方程是 20 世纪中叶发展起来的一个学科.系数为随机量的常微分方程和由随机过程驱动的微分系统,一般称为随机微分方程.又因为前者可以直接地处理为随机参数的常微分方程,所以,通常的随机微分方程常常专门指后者.由于快速变化的噪声可以用 Brown 运动建模,也由于这方面的理论研究成果已经很充分,处理形式也相对地简单,并且在实际中也更多地出现,所以人们更关注于以 Brown 运动为驱动的随机微分方程,研究它的基本性质,利用它来建模.例如,在金融系统、数量经济、控制系统、统计物理、系统生物学中都常见到这样的模型.理论的发展与应用的需要就形成了以 Ito 积分为核心的 Ito 随机分析的学科.另一方面,由于随机驱动中突变的发生,人们又引入了以 Poisson 点过程为驱动的随机分析,称为 Poisson 随机分析.这两种最典型的模型,相对地又具有许多使用方便的性质,特别地,鞅论的成果为此提供了有力的数学工具.

在统计物理中,人们是通过求解 Master 方程和 Fokker-Plank 方程得到系统状态的概率密度和状态转移的概率密度,由此了解系统的时间发展的.事实上,仅有这样的理解是片面的.要想了解随机系统的时间发展,更需要将状态按时间的发展作为随机试验的基本结果进行概率和统计研究,这就是随机过程的轨道.因为描写转移的统计概率规律的参数通常是不知道的,所以,只知道系统在固定时间的统计规律和状态转移的统计规律常常是不够的,这时就需要通过实验的结果统计地推定,即作统计估计.为此就需要用随机过程的一段足够长时间的现实,即轨道数据,来估值.这种用随机过程的一个轨道表示过程的特征参数的性质,就是用时间平均近似空间平均的性质,统计物理学家通常称之为遍历

性质.统计物理的基础之一就是假定系统的发展具有遍历性质.然而如果进一步追溯，什么样的模型能具有遍历性质呢？统计物理中忽视了这一方面的阐述与研究.而用 Markov 过程与随机微分方程的理论可以完美地回答这个问题.这个理论说明在很宽松的条件下，用随机过程的一个轨道表示过程的特征参数是完全可能的.由此可见，了解随机过程的精髓就要了解系统模型的状态的时间发展，即随机过程的轨道的研究.在这方面随机微分方程提供了一个强大的模型与有力的工具.

本书的内容安排，首先给出概率论扼要的回顾，作为初等概率论的补充，还简要地介绍了随机数的生成与 Gauss 系.我们以更多的篇幅阐明条件推理，由具体情形开始，逐步地将它一般化.这就是条件分布与条件期望.第 3 章的随机徘徊无论在理论上或在近似计算中都是基础性的内容，而后的鞅列简单介绍了鞅这个有力的概率论工具，并且分析了一些典型的例子，介绍了在应用中极为重要的选样定理.第 4 章集中讨论 Brown 运动，并且顺带介绍了 Markov 过程，这是在应用中最常见的一种模型.第 5 章是随机微积分的基础，主要是对 Brown 运动的 Ito 积分与 Ito 公式.为了了解其发展的历史过程与思想轨迹，我们不惜花较大的篇幅，对于被积分的对象从最简单的情形开始，由此看出积分的基本特性，发现其局限之处，构想先驱们用以突破局限性并扩展概念的思路.具体地说，就是对于 Ito 积分的被积函数，先从连续函数和平方可积的数值函数入手，然后对于由 Brown 运动的历史所决定的连续随机过程，继而对于由 Brown 运动的历史所决定的“平方可积”的随机过程定义平均收敛意义下的 Ito 积分，最后，由于涉及的 Ito 积分的 Ito 过程对于复合函数不能做到完全封闭，所以有必要引入所有的轨道都平方可积的，且由 Brown 运动的历史所决定的“局部平方可积”的随机过程，按照依概率收敛扩展定义的 Ito 积分.这种积分是首先由 Skorohod 在 20 世纪 60 年代初给出的，其后 Ito 学派将它简化为由 Brown 运动的历史所决定的“平方可积”的随机过程的 Ito 积分沿轨道的延伸.经过如此扩展的 Ito 积分所涉及的 Ito 过程对于光滑函数的复合具有封闭性.这就导致精致而完备的 Ito 公式，它是随机分析的核心所在.第 6 章是本书的中心论题，书中并没有给出解的存在唯一性的证明，但是需要强调的是解的存在唯一性的条件是用随机微分方程建模正确性的理论保证.在本章中，还对 Kalman-Bucy 滤波模型作了较为详细的阐述.第 7 章讨论随机微分方程解的 Markov 性质，以及这种扩散过程的分布密度所满足的 Master 方程和转移密度所满足的 Kolmogorov 方程，或 Fokker-Plank 方程.这些是物理学家的基本统计工具.顺便指出，一般的 Markov 过程模型及其转移密度的向前与向后两个基本方程，是由 Kolmogorov 在 20 世纪 30 年代首创的.在本章中阐述了在时间发展长久时，非退化扩散过程的转移密度的两歧性质，以及不变密度所满足的方程与遍历定理.在离散时间情形，这可以用来统计地估计 Markov 链(它是 Markov 过程在离散时间采样与离散状态近似下的模型)的平稳概率与转移参数.本章随后介绍的 Girsanov 定理与 Feyman-Kac 公式是随机分析中最基本的公式.它们与 Ito 公式一起被称为随机分析的三大支柱.由于它们

具有的理论深度需要较严密的随机过程理论基础才能剖析清楚，在本章中只是让读者有一个初步的了解与认知. 而进一步的了解还需要查阅随机过程、随机微分方程方面的理论教材或文献. 本章的最后一节介绍最佳停止，这个内容涉及随机过程、最佳控制、金融期权与微分方程活动边界等学科. 这里只是介绍了一个简单情形. 第 8 章给出得到随机微分方程一个样本轨道的数值模拟方法，以及逐步提高精度的途径，并且介绍了几个常用的近似模型. 第 9 章介绍金融中的随机微分方程模型. 以 20 世纪 90 年代后期获得 Nobel 经济奖的 Black-Scholes 在 1973 年引入的模型为主线展开，介绍了在此模型下得到欧式未定权益的定价公式的几种常见方法. 而后讨论作为其简单的近似模型：二叉模型. 对二叉模型具体地给出了欧式未定权益的定价和美式未定权益的定价函数的算法. 本章最后简单地讨论债券的随机利率，利率的随机模型和利率衍生证券. 第 10 章是最后一章，扼要地阐述了 Poisson 随机分析，包括以 Poisson 点过程为驱动的随机微分方程. 还介绍了常见的条件 Poisson 过程，包括自激点过程与带调制的 Poisson 过程. 这样就扩大了建模的库存. 本章后面介绍了在理论上研究随机过程的特征泛函，其重要性在于它与过程的有限维分布族是等价的，也就是由它可以决定过程的统计分布. 最后给出了模拟 Poisson 事件与非时齐 Poisson 事件的思路.

带 * 号的内容初次阅读时可以跳过.

随机微分方程的数学基础，对于非数学，甚至非概率论的读者颇为艰深. 要想撰写一本作为初等概率论的直接后继教材的尝试，笔者也颇感勉为其力. 但是，鉴于读者需要这样一本较为浅近的材料，笔者希望在有生之年奋力一试，以此奉献初学者，算是交个朋友. 此外，各位有什么建议，请不吝反馈. 笔者在此预先致谢.

最后请读者注意，一本书的前言有时很精炼地概括了章节中的主线，随着课文的进行反复地阅读品味将会加深体会. 此外，目录与名词索引能快速地帮助了解、浏览一些重要的概念. 请珍视利用. 笔者至今深深地记住年轻时老师的教诲：读一本书要学会阅读前言、目录、名词索引与符号表. 这样就不是简单地按一维的方式阅读，而是按二维的方式阅读.

龚光鲁　2007 年
glgong@math. tsinghua. edu. cn

符　号　表

$[a]$　不超过 a 的最大整数

$a \wedge b$　a 和 b 中小的一个

$a \vee b$　a 和 b 中大的一个

$b(t,x,y)$　Brown 运动的转移密度

B_t　Brown 运动

$\stackrel{\text{def}}{=}$　表示定义为

$\boldsymbol{B}_s = \{B_u: u \leqslant s\}, B_{\leqslant t}$　Brown 运动在时间 s 前的集合

$B(n,p)$　二项分布

C_0^2　两次连续可微的紧支集函数类

$\mathrm{cov}(\xi,\eta)$　随机变量 ξ,η 的协方差

$E\xi$　随机变量 ξ 的期望

$E(\xi|\eta), E(\xi|\eta_t, t \geqslant 0), E(\xi|\mathscr{F})$　随机变量 ξ 对 η，对 $\{\eta_t, t \geqslant 0\}$，对 $\mathscr{F}$ 的条件期望

Exp_λ　参数为 λ 的指数分布

$f_\xi(x|\eta=y), f_{\xi|\eta}(x|y), f_{\boldsymbol{\xi}}(\boldsymbol{x}|\boldsymbol{y}), f_{\boldsymbol{\xi}}(\boldsymbol{x}|\boldsymbol{\eta}=\boldsymbol{y})$　随机变量 ξ 对于 $\eta=y$ 的条件密度

$f(\xi_t) \circ \mathrm{d}B_t$　Stratonovich 微分

$f * g$　卷积

$\mathscr{F}$　事件体，σ 代数

$F^{-1*}(z) = \inf(x: F(x) > z)$　$F(x)$的广义逆函数

$\Gamma(x)$　Γ 函数，伽玛函数

$\Gamma(\theta,\alpha)$　Γ 分布，伽玛分布

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

$L^2(\Omega)$方差有限的随机变量全体

$L^2(\bar{\Phi}(\xi))\{\xi_t: t \in I\}$的非线性均方信息空间

$\mathscr{L}^1 \stackrel{\text{def}}{=} \left\{(B_t)\text{ 可知的随机过程 } \Phi_t\text{，且}\int_0^T E\,|\,\Phi_t(\omega)\,|\,\mathrm{d}t < +\infty\right\}$

$\mathscr{L}^2 \stackrel{\text{def}}{=} \left\{(B_t)\text{ 可知的随机过程 } \Phi_t\text{，且}\int_0^T E\,|\,\Phi_t(\omega)\,|^2\,\mathrm{d}t < +\infty\right\}$

$\mathcal{L}^{1,\mathrm{loc}} \stackrel{\mathrm{def}}{=} \left\{ (B_t) \text{ 可知的随机过程 } \Psi_t, \text{且 } P\left(\omega: \int_0^T |\Psi_t(\omega)| \mathrm{d}t < +\infty\right) = 1 \right\}.$

$\mathcal{L}^{2,\mathrm{loc}} \stackrel{\mathrm{def}}{=} \left\{ (B_t) \text{ 可知的随机过程 } \Phi_t, \text{且 } P\left(\omega: \int_0^T |\Phi_t(\omega)|^2 \mathrm{d}t < +\infty\right) = 1 \right\}.$

$\bar{L}(\xi)$：$\{\xi_t: t \in I\}$ 的线性闭包

$\log N(\mu,\sigma^2)$ 对数正态分布

$M(z) = E\mathrm{e}^{z\xi}$ 随机变量 ξ 的矩母函数

$N(\mu,\sigma^2)$，$N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ 正态分布

$\mathrm{Poisson}_\lambda$ 参数为 λ 的 Poisson 分布

$\mathrm{Pro}_V(x)$，$\mathrm{Proj}_{\bar{L}(\xi)}\eta$ x 在空间 V 上的投影，η 在空间 $L(\xi)$ 上的投影

T_a 首达 a 的时刻

$\boldsymbol{A}^{\mathrm{T}}$ 矩阵 $\boldsymbol{A}$ 的转置

$\rho_{\xi\eta}$ 随机变量 ξ 与 η 的相关系数

$\xi_n \xrightarrow{p} \xi$ 随机变量序列 $\{\xi_n\}$ 依概率收敛到随机变量 ξ

$\xi_n \xrightarrow{L^2} \xi$ 随机变量序列 $\{\xi_n\}$ 均方收敛到随机变量 ξ

$\xi_n \xrightarrow{\mathrm{a.e.}} \xi$ 随机变量序列 $\{\xi_n\}$ 概率为 1 地收敛到随机变量 ξ

$\xi_n \xrightarrow{d} \xi$ 随机变量序列 $\{\xi_n\}$ 依分布收敛到随机变量 ξ

$\xi_n \xrightarrow{w} \xi$ 随机变量序列 $\{\xi_n\}$ 弱收敛到随机变量 ξ，即依分布收敛

$(\xi,\eta) \stackrel{\mathrm{def}}{=} E(\xi\eta)$

$((\Phi,\Psi)) \stackrel{\mathrm{def}}{=} \int_0^T E(\Phi_t\Psi_t)\mathrm{d}t$

$|||\Phi|||^2 \stackrel{\mathrm{def}}{=} \int_0^T E|\Phi_t|^2 \mathrm{d}t$

$\bar{\Phi}(\eta)$：$\{\eta_t: t \in I\}$ 的非线性信息空间

$\varphi(t) = E\mathrm{e}^{\mathrm{i}t\xi}$ 随机变量 ξ 的特征函数

$\boldsymbol{\Sigma}_{\boldsymbol{\xi}} = (\mathrm{cov}(\xi_i,\xi_j))_{i,j\leqslant d}$ d 维随机向量 $\boldsymbol{\xi} = (\xi_1,\xi_2,\cdots,\xi_d)^{\mathrm{T}}$ 的(协)方差阵

$U[a,b]$ $[a,b]$ 区间上的均匀分布

$\mathrm{var}\xi$ 随机变量 ξ 的方差

$\boldsymbol{\eta}_t \stackrel{\mathrm{def}}{=} \{\eta_s: s \leqslant t\}$

ω 基本事件，样本点

Ω 样本空间

目　录

第 1 章　概率论备要与随机数

1.1　概率论备要

本节对于概率论的基本概念作一个简短的复习.

1.1.1　概率公理系统

一个随机试验可能出现的一个结果,称为一个**基本事件**,或**样本点**,记为 ω. 全体基本事件的集合记为 Ω,称为必然事件,或**样本空间**. Ω 的某些子集组成的类 $\mathscr{F}$,称为一个**事件体**,或 **σ 代数**,如果它满足:$\Omega\in\mathscr{F}$,且

$$A,A_n \in \mathscr{F}(n\geqslant 1)\Rightarrow \bigcup_{n=1}^{+\infty} A_n,\ \bigcap_{n=1}^{+\infty} A_n,A^c \overset{\text{def}}{=\!=} \Omega-A \in \mathscr{F}.$$

$\mathscr{F}$ 中的集合称为**随机事件**,在直观上可以理解为可以描述其概率的事情,或有概率的事情.

在 $\mathscr{F}$ 上定义的非负函数 $P(A)$(任意 $A\in\mathscr{F}$),称为**概率**,如果它满足:

(1) $P(\Omega)=1$.

(2) 对于任意 $A_i\in\mathscr{F}(i\geqslant 1)$,只要 $A_i\cap A_j\neq\varnothing(i\neq j)$,其中 $\varnothing$ 表示没有基本事件的空集,就有

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i).$$

这里 $P(A)$ 表示事件 A 发生的概率.

注　当样本空间与事件体都确定以后,Kolmogorov 公理系统仍旧允许容纳不止一个“取概率运算”$\{P(A):A\in\mathscr{F}\}$. 初学者很难理解这一点,因为一般初学者常直观地认为,当样本空间与事件体都确定以后,概率就自然地被确定了. 事实上,局限于这种想法会削弱概率公理系统的机变与处理能力. 在同一个样本空间与同一个事件体上,可以存在不同的取概率运算. 这在有的时候表现为概率间的变换,这是概率论在理论与应用中常用的技巧,在第 9 章的金融数学模型中,由 Black-Scholes 模型到中性概率模型的变化,就是一个比较自然的概率变换的例子.

1.1.2 随机变量

一个随机地取实数值的量 $\xi=\xi(\omega)(\omega\in\Omega)$，称为**随机变量**，如果对于任意实数 x，样本点的集合$\{\omega: \xi(\omega)\leqslant x\}$都是一个随机事件.

随机变量的定义依赖于给定的事件体. 而实值函数

$$F(x) \stackrel{\text{def}}{=} P(\xi \leqslant x)$$

称为随机变量 ξ 的**分布函数**.

离散随机变量 ξ 的**概率函数**(或**概率分布**)为

$$f(x) \stackrel{\text{def}}{=} P(\xi = x) \quad (x = x_1, x_2, \cdots).$$

对于实值函数 $g(x)$，随机变量 $g(\xi)$的期望为

$$Eg(\xi) = \sum_i g(x_i) f(x_i).$$

连续型随机变量 ξ 的**分布密度**为 $f(x)$，其分布函数为 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$. 而实值函数 $g(\xi)$ 的期望为

$$Eg(\xi) = \int_{-\infty}^{+\infty} g(x) f(x)\mathrm{d}x.$$

若 ξ 为非负随机变量，则还有

$$E\xi = \int_0^{+\infty} P(\xi > a)\mathrm{d}a$$

随机变量 ξ 的**方差**定义为

$$\operatorname{var}\xi = E(\xi - E\xi)^2 = E\xi^2 - (E\xi)^2.$$

矩母函数为

$$M(z) = E\mathrm{e}^{z\xi} \quad (\text{如果它有限}),$$

它包含了任意阶的矩 $E\xi^k$(**k 阶矩**)的信息，而且 ξ 的分布也可由矩母函数唯一确定. 如果矩母函数不是有限的，则可改用**特征函数**

$$\varphi(t) = E\mathrm{e}^{\mathrm{i}t\xi}$$

1.1.3 随机向量

d 维随机向量$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_d \end{pmatrix}$的分布函数为 $F(\boldsymbol{x})=P(\boldsymbol{\xi}\leqslant\boldsymbol{x})$，这里$\boldsymbol{\xi}\leqslant\boldsymbol{x}$ 的含义为

$$\xi_1 \leqslant x_1, \quad \xi_2 < x_2, \quad \cdots, \quad \xi_d \leqslant x_d,$$

其数学期望和(协)方差阵，分别为

$$E\boldsymbol{\xi}=\begin{pmatrix}E\xi_1\\E\xi_2\\\vdots\\E\xi_d\end{pmatrix},\quad \boldsymbol{\Sigma}_{\boldsymbol{\xi}}=(\mathrm{cov}(\xi_i,\xi_j))_{i,j\leqslant d}.$$

其中

$$\mathrm{cov}(\xi,\eta)=E(\xi\eta)-E\xi E\eta$$

是(ξ,η)的双线性函数,称为**协方差**. **相关系数**为

$$\rho_{\xi\eta}=\frac{\mathrm{cov}(\xi,\eta)}{\sqrt{\mathrm{var}\xi}\ \sqrt{\mathrm{var}\eta}}.$$

$\boldsymbol{\xi}$的矩母函数为

$$M(\boldsymbol{z})=E\mathrm{e}^{\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\xi}},$$

其中$\boldsymbol{z}=(z_1,z_2,\cdots,z_d)^{\mathrm{T}}$,而$(\quad)^{\mathrm{T}}$表示向量或矩阵的转置运算.

特征函数为

$$\Phi(\boldsymbol{t})=E\mathrm{e}^{\mathrm{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\xi}},$$

其中$\boldsymbol{t}=(t_1,t_2,\cdots,t_d)^{\mathrm{T}}$.

随机变量组$\{\xi_1,\xi_2,\cdots,\xi_n\}$称为**独立**,如果它们满足条件

$$P(\xi_1\leqslant x_1,\xi_2\leqslant x_2,\cdots,\xi_n\leqslant x_n)=P(\xi_1\leqslant x_1)P(\xi_2\leqslant x_2)\cdots P(\xi_n\leqslant x_n).$$

两个随机向量$\boldsymbol{\xi},\boldsymbol{\eta}$称为独立,如果对于规则的集合(即1.1.4节中将阐明的Borel集合)A,B,事件$\boldsymbol{\xi}\in A$与$\boldsymbol{\eta}\in B$恒为独立,这等价于

对于任意的f,g恒有$E[f(\boldsymbol{\xi})g(\boldsymbol{\eta})]=Ef(\boldsymbol{\xi})Eg(\boldsymbol{\eta})$.

设随机变量ξ,η独立且分别具有密度$f(x),g(x)$,则其和$\xi+\eta$具有**卷积密度**

$$(f*g)(x)\xlongequal{\mathrm{def}}\int f(x-u)g(u)\mathrm{d}u=(g*f)(x).$$

如果当$x<0$时,$f(x)=g(x)=0$,则

$$(f*g)(x)=\int_0^x f(x-u)g(u)\mathrm{d}u.$$

对于ξ为随机变量,恒有**Chebyshev不等式**

$$P(|\xi-E\xi|\geqslant\delta)\leqslant\frac{\mathrm{var}\xi}{\delta^2}.$$

1.1.4 基本极限定理

如果随机变量序列$\{\xi_n\}$与随机变量ξ间满足:对于任意固定的$\varepsilon>0$,恒有

$$P(|\xi_n-\xi|\geqslant\varepsilon)\xrightarrow{n\to+\infty}0,$$

则称随机变量序列$\{\xi_n\}$**依概率收敛**到随机变量ξ,记为$\xi_n\xrightarrow{p}\xi$,其含义为,对于一个任意

给定的误差 ε，在忽略一个任意小概率的情形下，ξ_n 可以近似 ξ.

对于连续函数 $g(\boldsymbol{x})$，我们恒有

$$\boldsymbol{\xi}^{(n)} \xrightarrow{p} \boldsymbol{\xi} \quad \Rightarrow g(\boldsymbol{\xi}^{(n)}) \xrightarrow{p} g(\boldsymbol{\xi}).$$

常用的**大数定律**是（称为 **Khinchine 大数定律**）：设当 $\{\xi_n\}$ 为独立同分布的随机变量序列，且

$$E\xi_n = \mu(n = 1,2,\cdots), \quad 则 \quad \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{n} \xrightarrow{p} \mu.$$

如果随机变量序列 $\{\xi_n\}$ 与随机变量 ξ 之间满足

$$E \mid \xi_n - \xi \mid^2 \to 0,$$

则称 $\{\xi_n\}$ **均方收敛**到 ξ，记为 $\xi_n \xrightarrow{L^2} \xi$.

由 Chebyshev 不等式立刻可以得到，均方收敛一定能推出依概率收敛.

如果随机变量序列 $\{\xi_n\}$ 与随机变量 ξ 之间满足

$$P(\xi_n \xrightarrow{n \to +\infty} \xi) = 1,$$

即"$\xi_n \to \xi$"是一个概率为 1 的事件，则称 $\{\xi_n\}$ **概率为 1 地收敛**到 ξ，记为 $\xi_n \xrightarrow{\text{a. e.}} \xi$. 这里 a. e. 是 almost everywhere 的缩写. 实函数分析指出：概率为 1 地收敛一定可以推出依概率收敛.

设 ξ 为连续型随机变量，其分布函数为 $F(x)$，如果随机变量序列 $\{\xi_n\}$ 的分布函数 $\{F_n(x)\}$ 收敛到 $F(x)$，则称 $\{\xi_n\}$ **依分布收敛**到 ξ，记为 $\xi_n \xrightarrow{d} \xi$.

对于非连续型的随机变量 ξ，如果在随机变量 ξ 的分布函数 $F(x)$ 的所有的连续点 x 上有：随机变量序列 $\{\xi_n\}$ 的分布函数 $F_n(x)$ 收敛到 $F(x)$，则也称 $\{\xi_n\}$ **依分布收敛**到 ξ. 依分布收敛也称为**弱收敛**，所以也记为 $\xi_n \xrightarrow{w} \xi$.

从依概率收敛一定能推出依分布收敛. 反之，从依分布收敛却不能推出依概率收敛，除非极限随机变量为一个常数.

概率理论证明了以下的 3 个叙述彼此等价

(1) $\xi_n \xrightarrow{d} \xi$；

(2) 特征函数 $\varphi_{\xi_n}(\theta) \stackrel{\text{def}}{=} E\mathrm{e}^{i\theta\xi_n} \to$ 特征函数 $\varphi_{\xi}(\theta) \stackrel{\text{def}}{=} E\mathrm{e}^{i\theta\xi}$

或

(2)′（当矩母函数有限时）

矩母函数 $M_{\xi_n}(z) \stackrel{\text{def}}{=} E\mathrm{e}^{z\xi_n} \to$ 矩母函数 $M_{\xi}(z) \stackrel{\text{def}}{=} E\mathrm{e}^{z\xi}$；

(3) 对于任意有界连续函数 g 有 $Eg(\xi_n) \to Eg(\xi)$.

常用的**中心极限定理**为（称为 Levy-Lindeberg 中心极限定理）

若随机变量序列$\{\xi_n\}$为独立同分布，$E\xi_n=\mu$ 且 $\mathrm{var}\xi_n=\sigma^2(n=1,2,\cdots)$，则

$$\frac{\xi_1+\xi_2+\cdots+\xi_n-n\mu}{\sqrt{n}\,\sigma}\xrightarrow{d}N(0,1).$$

对于 d 维随机变量，也有相应的按分布(即按联合分布)收敛的概念及中心极限定理.

1.1.5 常见的实函数，Borel 事件体与 Borel 函数

常见的实函数有

(1) 初等函数：多项式函数，三角函数，指数函数，对数函数及它们的组合与复合.

(2) 常见的非初等的连续可微函数有**Γ 函数**(Gamma 函数)

$$\Gamma(x)=\int_0^{+\infty}u^{x-1}\mathrm{e}^{-u}\mathrm{d}u\quad(x>0).$$

① $\Gamma(1)=1$.

② 用分部积分，容易证明如下的递推公式：

$\Gamma(x+1)=x\Gamma(x)$ (特别地，$\Gamma(n+1)=n!$，故它是阶乘的推广).

③ $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.

③ 的证明 比较两种方法计算下述积分：$I=\int_0^{+\infty}\mathrm{e}^{-x^2}\mathrm{d}x$. 首先，用变量替换 $x^2=u$ 得到 $I=\frac{1}{2}\int_0^{+\infty}u^{-\frac{1}{2}}\mathrm{e}^{-u}\mathrm{d}u=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$. 另一方面，通过利用对称性与取极坐标得到

$$I^2=\int_0^{+\infty}\mathrm{e}^{-x^2}\mathrm{d}x\cdot\int_0^{+\infty}\mathrm{e}^{-y^2}\mathrm{d}y=\iint\limits_{\text{第一象限}}\mathrm{e}^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y$$

$$=\int_0^{\frac{\pi}{2}}\int_0^{+\infty}\mathrm{e}^{-r^2}r\mathrm{d}r=\frac{\pi}{4}\int_0^{+\infty}\mathrm{e}^{-u}\mathrm{d}u=\frac{\pi}{4}.$$

将两者比较便得到 $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.

(3) 常见的重要不连续函数：集合 A 的**示性函数**(indicator)

$$I_A(x)=\begin{cases}1, & x\in A,\\ 0, & x\notin A.\end{cases}$$

特例：$A=[a,b]$，$A=\{x_0\}$. 示性函数是最重要的不连续函数，它的重要性差不多等同于初等函数.

包含所有开区间的最小的事件体中的集合，称为 **Borel 事件体**.

包含所有 d 维开矩形的最小的事件体中的集合，称为 **d 维 Borel 事件体**.

包含所有连续函数，且对于(取函数的逐点)极限运算封闭和对线性运算封闭的最小函数类，称为 **Borel 函数类**，其中的函数，称为 **Borel 函数**.

1.1.6 常见的概率分布

分布名称	参数	分布列或概率密度	数学期望 $E\xi$	方差 $\mathrm{var}\xi$	特征函数 Ee^{itX}
两点分布 (Bernoulli 分布)	$0<p<1$	$f(k)=\begin{cases}p, & k=1\\ q, & k=0\end{cases}\quad (q=1-p)$	p	pq	$q+pe^{it}$
二项分布 $B(n,p)$	$n\geqslant 1$ $0<p<1$	$f(k)=C_n^k p^k q^{n-k}$ $(k=0,1,\cdots,n,q=1-p)$	np	npq	$(q+pe^{it})^n$
几何分布 Geo_p	$0<p<1$	$f(k)=pq^{n-1}\quad (k=1,2,\cdots,q=1-p)$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^{it}}{1-qe^{it}}$
超几何分布	N,M,n $(n\leqslant M)$	$f(k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}\quad (k=0,1,\cdots,n)$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$	略
Poisson 分布 $\mathrm{Poisson}_\lambda$	$\lambda>0$	$f(k)=\frac{\lambda^k e^{-\lambda}}{k!}\quad (k=0,1,\cdots)$	λ	λ	$e^{\lambda(e^{it}-1)}$
均匀分布 $U[a,b]$	$a<b$	$f(x)=\frac{1}{b-a}I_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{itb}-e^{ita}}{(b-a)i}$
正态分布 $N(\mu,\sigma^2)$	μ $\sigma>0$	$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{it\mu-\frac{1}{2}t^2\sigma^2}$
对数正态分布 $\log N(\mu,\sigma^2)$	μ $\sigma>0$	$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma x}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}I_{(0,+\infty)}(x)$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$	略
指数分布 Exp_λ	$\lambda>0$	$f(x)=\lambda e^{-\lambda x}I_{[0,+\infty)}(x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\left(1-\frac{it}{\lambda}\right)^{-1}$
Γ 分布 $\Gamma(\theta,\alpha)$ 分布	$\theta>0$ $\alpha>0$	$f(x)=\frac{1}{\Gamma(\alpha)\theta}\left(\frac{x}{\theta}\right)^{\alpha-1}e^{-\frac{x}{\theta}}I_{[0,+\infty)}(x)$	$\alpha\theta$	$\alpha\theta^2$	$(1-i\theta t)^{-r}$
Beta 分布 $B(\alpha,\beta)$ 分布	$\theta>0$ $\alpha>0$	$f(x)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}x^{\alpha-1}(1-x)^{\beta-1}I_{[0,1]}(x)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	略

1.2 随机数与随机模拟

当前，在建模中应用随机方法的重要手段是随机模拟. 本节介绍作为随机模拟的基础的随机数的生成方法.

1.2.1 生成随机数的逆函数方法

定义 1.1 $[0,1]$上均匀分布的随机变量的 n 个独立的样本,称为 n 个**均匀随机数**,简称**随机数**.

在 Microsoft Excel 菜单中,点击:

插入→函数→全部→RAND,

则可以由计算机生成一个数,人们的长期实践经验证实,将它视为$[0,1]$上均匀分布的随机变量的取值是可行的.

定义 1.2 若随机变量 X 的分布函数为 $F(x)$,则 X 的一个样本值称为一个 ***F*-随机数**. 在 $F(x)$有密度函数 $f(x)$时,也称为 ***f*-随机数**.

命题 1.1(逆函数方法) 如果分布函数 $F(x)$严格单调,u 是一个均匀随机数,则 $F^{-1}(u)$(其中 F^{-1}为 F 的逆函数)是一个 ***F*-随机数**.

证明 设 U 是一个$[0,1]$上的均匀随机变量,那么

$$P(F^{-1}(U)\leqslant x)=P(U\leqslant F(x))=F(x).$$

注 如果 $F(x)$并非严格单调,我们可以定义函数 $F(x)$的如下的"广义的逆函数":

$$F^{-1*}(z)=\inf\{x: F(x)>z\}.$$

如果 u 是一个均匀随机数,那么,可以证明 $F^{-1*}(u)$是一个 F-随机数.

使用计算机可以生成各种分布的随机数,这样,就可以在计算机上作重复的随机试验,从而可以得到许多模拟数据.模拟数据能为随机建模与对随机模型的运行作出合理的分析,提供重要的参考作用.

例 1.1(离散分布随机数) 设离散随机变量 X 的分布函数为 $F(x)$,概率函数为

$$f(x)=P(X=x),\quad 满足\ p_i=f(x_i)>0,\quad \sum_i p_i=1.$$

我们将$[0,1]$分为一些互不相交的子区间,使第 i 个子区间 J_i 的长度为 p_i.任取一个均匀随机数 u,我们定义一个由它生出 Z 的如下方法:

$$若\ u\in J_i,\quad 令\ z=x_i(意思是:如果\ u\ 落入\ J_i,就取\ z=x_i),$$

那么,Z 是一个 F-随机数.

例 1.2(集合$\{1,2,\cdots,n\}$上的离散均匀随机数) 设 u 是均匀随机数,那么 $z=[nu]+1$ 是集合$\{1,2,\cdots,n\}$上的离散均匀随机数,其中记号$[a]$表示正数 a 的整数部分.事实上

$$z=j\Leftrightarrow u\in\left[\frac{j-1}{n},\frac{j}{n}\right)\Leftrightarrow nu\in[j-1,j)\Leftrightarrow j=[nu]+1.$$

例 1.3(几何随机数) 设随机变量 X 服从几何分布

$$f(k)\overset{\text{def}}{=}P(X=k)=q^{k-1}p\quad(k\geqslant 1,q=1-p).$$

我们注意,对于 $U\sim U[0,1]$有

$$P(1-q^{k-1}\leqslant U<1-q^k)=P(X=k).$$

然而 $1-q^{X-1}\leqslant U<1-q^X$ 等价于 $X\cdot\ln q<\ln(1-U)<(X-1)\ln q$. 后者又等价于 $X=\left[\frac{\ln(1-U)}{\ln q}\right]+1$，其中方括号表示取整运算. 因此，对于均匀随机数 u，

$$k\overset{\text{def}}{=}\left[\frac{\ln(1-u)}{\ln q}\right]+1$$

就是参数为 $p=1-q$ 的几何随机数.

注意，当 U 是 $U[0,1]$上随机变量时，$1-U$ 也是 $U[0,1]$随机变量. 故而，对于均匀随机数 u，$x=\left[\frac{\ln u}{\ln q}\right]+1$ 也是一个几何随机数. 例如，$p=0.4$ 时有

u	0.124	0.450	0.467	0.286	0.817
$x=\left[\frac{\ln u}{\ln q}\right]+1$	5	2	2	3	1

例 1.4(指数随机数)　参数为 λ 的指数分布的分布函数为 $F(x)=1-e^{-\lambda x}$，此时有 $F^{-1}(y)=-\frac{1}{\lambda}\ln(1-y)$. 任取一个均匀随机数 u，那么 $z=-\frac{1}{\lambda}\ln(1-u)$就是一个指数随机数. 类似地，对于均匀随机数 u，$-\frac{1}{\lambda}\ln u$ 也是一个指数随机数.

例 1.5(正态随机数)　如果 u 是均匀随机数，那么，$\Phi^{-1}(u)$是标准正态随机数，其中 $\Phi(x)$是标准正态分布函数. 而 $\sigma\Phi^{-1}(u)+\mu$ 就是参数为(μ,σ^2)的正态随机数.

例 1.6　自来水公司给某个企业每月分配 400000t 水，超标用水每 1000t 罚款 100 元. 假定公司每月实际用水 Wt，服从正态分布 $N(350000,35000^2)$. 为此公司决定用逆函数方法，模拟下一个季度的 W 值. 如果 $u_1=0.8643$，$u_2=0.3821$，$u_3=0.9452$ 是用来得到 W 值的均匀随机数. 我们有

$$\Phi^{-1}(u_1)=1,\quad \Phi^{-1}(u_2)=-0.3,\quad \Phi^{-1}(u_3)=1.6.$$

于是

$$w_i=356000+25000\Phi^{-1}(u_i)\quad (i=1,2,3)$$

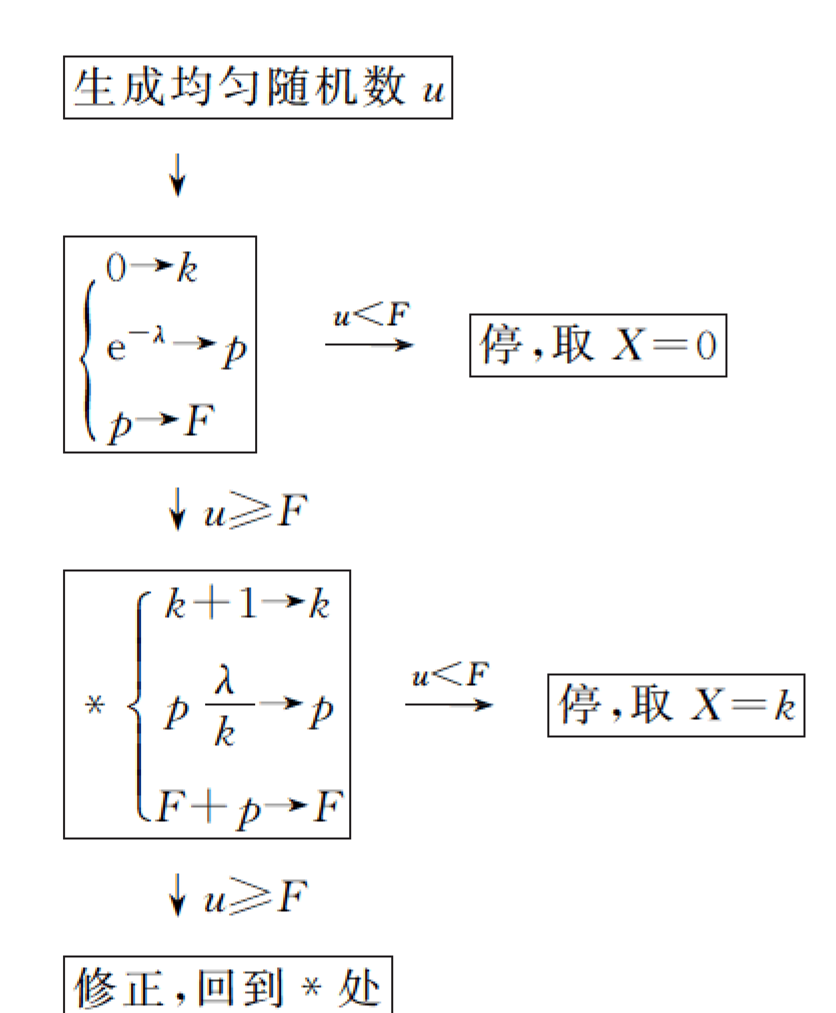
是 3 个 $N(350000,35000^2)$随机数：388500，339500，406000. 这样，下一个季度罚款的模拟值为 $100\times\frac{0+0+6000}{1000}=600$ 元.

Poisson 随机数可以按例 1.1 的方法生成. 但是，下述方法的效率更高.

例 1.7(Poisson 随机数)　对于 Poisson 随机变量 X，我们有递推公式

$$P(X=0)=e^{-\lambda},\quad P(X=k)=\frac{\lambda}{k}P(X=k-1)\quad (k\geqslant 1).$$

生成 Poisson 随机数可用下述**向上搜索流程**(k,p,F 是三个"变量")：



例如,由均匀随机数 $u=0.628$,求一个 Poisson_3 随机数.我们用上述流程:

$$k=0,\quad p=e^{-3}=0.05,\quad F=0.05<u,\quad k=1,\quad p=0.15,\quad F=0.2<u,$$
$$k=2,\quad p=0.224,\quad F=0.424<u,\quad k=3,\quad p=0.224,\quad F=0.648>u.$$

所以,一个 Poisson_3 随机数是 3.

对于较大均值的 Poisson 随机数,则可以使用**向下搜索流程**.

注 另一种生成Poisson随机数的更为常用的方法是:生成相互独立的均匀随机数 u_1, $u_2,\cdots$,取非负整数 m 使 $\prod_{k=1}^{m+1}u_k<e^{-\lambda}\leqslant\prod_{k=1}^{m}u_k$(即 $m=\min\{k\geqslant 1: u_1,\cdots,u_k<e^{-\lambda}\}-1$),则 m 是 Poisson 随机数.

证明 设 U_k 为相互独立的 $U[0,1]$随机变量,则 $T_k\overset{\text{def}}{=\!=}-\ln U_k$ 是相互独立的参数为 1 的指数随机变量. 令 $\tau_k\overset{\text{def}}{=\!=}T_1+T_2+\cdots+T_k$,那么,$N_t\overset{\text{def}}{=\!=}\sup\{k: \tau_k\leqslant t\}$是强度为 1 的 Poisson 过程. 于是

$$N_\lambda=m\ \text{等价于}\ \tau_m\leqslant\lambda,\quad \tau_{m+1}>\lambda,$$

即

$$-\ln\left(\prod_{k=1}^{m}U_k\right)<\lambda\leqslant-\ln\left(\prod_{k=1}^{m+1}u_k\right),$$

也就是

$$\prod_{k=1}^{m+1}U_k<e^{-\lambda}\leqslant\prod_{k=1}^{m}U_k.$$

例 1.8 保险公司为某市冬季 4 个月扫雪损失的理赔设置了 10000 元的免赔额. 假定每月损失是均值为 15000,标准差为 2000 的正态随机变量. 假定得到了 4 个均匀随机数

$$0.5398,\quad 0.1151,\quad 0.0013,\quad 0.7881$$

公司对于 4 个月扫雪损失的理赔额作了模拟计算. 用 Microsoft Excel 由均匀随机数 0.5398, 0.1151, 0.0013, 0.7881 分别得到的均值为 15000, 标准差为 2000 的正态随机数是

$$15200,\quad 12600,\quad 9000,\quad 16600.$$

因为只理赔超出免赔额的损失, 所以实际理赔的只有 3 份, 其理赔额分别为

$$5200,\quad 2600,\quad 6600.$$

于是公司对于 4 个月理赔总额的一个模拟值是 5200+2600+6600=14400.

例 1.9(混合分布随机数)　设

$$0=t_0<t_1<\cdots<t_n=1,\quad t_i-t_{i-1}=p_i(i\leqslant n),\quad F(x)=\sum_{i=1}^{n}p_iF_i(x),$$

$F_i(x)(i\leqslant n)$是分布函数. 又 u 是均匀随机数, 若 $t_{i-1}<u\leqslant t_i$, 则我们令

$$z=F_i^{-1}\left(\frac{u-t_{i-1}}{p_i}\right),$$

那么 z 是 F-随机数.

证明　对于[0,1]上的均匀随机变量 U, 定义随机变量

$$Z=F_i^{-1}\left(\frac{U-t_{i-1}}{p_i}\right)\quad (若\ t_{i-1}<U\leqslant t_i),$$

于是

$$\begin{aligned}P(Z\leqslant x)&=P\left(\sum_{i=1}^{n}F_i^{-1}\left(\frac{U-t_{i-1}}{p_i}\right)I_{(t_{i-1},t_i]}(U)\leqslant x\right)\\&=\sum_{i=1}^{n}P\left(F_i^{-1}\left(\frac{U-t_{i-1}}{p_i}\right)I_{(t_{i-1},t_i]}(U)\leqslant x,U\in(t_{i-1},t_i]\right)\\&=\sum_{i=1}^{n}P(t_{i-1}<U\leqslant t_{i-1}+p_iF_i(x))=\sum_{i=1}^{n}p_iF_i(x).\end{aligned}$$

即在计算时如果取到的均匀随机数 u 落在第 i 个区间$(t_{i-1},t_i]$中, 则取

$$z=F_i^{-1}\left(\frac{u-t_{i-1}}{p_i}\right),$$

这样得到的 z, 就是 $\sum_{i=1}^{n}p_iF_i$ 混合分布随机数.

混合分布最典型例子是**混合指数分布**, 即 $F_i(x)=1-e^{-\lambda_i x}$ 的情形, 此时

$$z=-\frac{1}{\lambda_i}\ln\frac{t_i-u}{p_i}\quad (t_{i-1}<u\leqslant t_i).$$

例 1.10(Gamma 随机数与 Beta 随机数)　n 个独立的均值为 θ 的指数随机数的和是参数为 n,θ 的 Gamma 随机数($\Gamma(n,\theta)$随机数). 若 y_1,y_2 分别是独立的 $\Gamma(n,1),\Gamma(m,1)$随机

数，那么，$\dfrac{y_1}{y_1+y_2}$是参数为 n,m 的 Beta 随机数(Beta(n,m)随机数).

1.2.2 生成随机数的 Von Neumann 取舍原则

给定分布密度 $f(x)$，**取舍原则**(rejection principle)提供了生成 f-随机数的一个简捷方法. 在生成 f-随机数时，逆函数方法往往需要很大的计算量. 而取舍原则的做法是：取一个参考分布密度函数 $g_0(x)$，使它满足：

(1) g_0-随机数容易生成. 如 $g_0(x)$为正态密度函数(双侧分布)，$U[a,b]$密度函数(有限区间分布)，指数密度函数(单侧分布)，及它们的混合密度函数(多峰分布).

(2) $g_0(x)$与 $f(x)$的取值范围差不多(不必相同)，且存在 C，使 $f(x)\leqslant Cg_0(x)$.

取舍原则的数学根据是以下的命题.

命题 1.2 设随机变量 Y 具有密度函数 $g_0(x)$，而随机变量 $U\sim U[0,1]$，且与 Y 独立，则

$$P\left(Y\leqslant x \mid \frac{f(Y)}{Cg_0(Y)}\geqslant U\right)=\int_{-\infty}^{x} f(v)\mathrm{d}v.$$

证明 对 Y 的取值用全概率公式$\left(P(A)=\int P(A\mid Y=x)g_0(x)\mathrm{d}x\right)$，可以得到

$$左=\frac{P\left(Y\leqslant x,\dfrac{f(Y)}{Cg_0(Y)}\geqslant U\right)}{P\left(\dfrac{f(Y)}{Cg_0(Y)}\geqslant U\right)}=\frac{\int_{-\infty}^{x}P\left(U\leqslant\dfrac{f(y)}{Cg_0(y)}\right)g_0(y)\mathrm{d}y}{\int_{-\infty}^{+\infty}P\left(U\leqslant\dfrac{f(y)}{Cg_0(y)}\right)g_0(y)\mathrm{d}y}=\frac{\int_{-\infty}^{x}\dfrac{1}{C}f(y)\mathrm{d}y}{\int_{-\infty}^{+\infty}\dfrac{1}{C}f(y)\mathrm{d}y}=右.$$

取舍原则的具体作法是：

(1) 独立地生成 n 个独立的 $g_0(x)$随机数 $v_1,v_2,\cdots,v_n$ 与 n 个与之独立的 $U[0,1]$随机数 $u_1,u_2,\cdots,u_n$.

(2) 对于(1)中生成的随机数按以下原则进行取舍：对 $i=1,2,\cdots$，如果有$\dfrac{f(v_i)}{Cg_0(v_i)}\geqslant u_i$，就保留 v_i，否则就舍弃 v_i.

由命题 1.2，所有保留下来的那些 v_i 就成为一系列独立 f-随机数.

推论 1.3 取舍原则可以改良为：如果 $f(x)=\gamma h(x)$，只要存在 C，使 $h(x)\leqslant Cg_0(x)$，那么，我们可以在取舍原则中用 $h(x)$代替 $f(x)$，得到 f-随机数. 其具体做法为：独立地生成 n 个独立的 g_0-随机数 $v_1,v_2,\cdots,v_n$ 与 n 个与之独立的均匀随机数 $u_1,u_2,\cdots,u_n$，如果$\dfrac{h(v_i)}{Cg_0(v_i)}\geqslant u_i$，则保留 v_i，否则舍弃 v_i，那么，所有保留下的 v_i 是相互独立的 f-随机数.

证明 只需注意到 $f(x)\leqslant\gamma Cg_0(x)$及$\dfrac{f(x)}{\gamma Cg_0(x)}=\dfrac{h(x)}{Cg_0(x)}$即可.

注1　推论1.3说明,我们不必计算 γ(即可以忽视计算量很大的一个常数因子).这样使取舍原则简单易行,可以适用于采集相当复杂的分布的随机数.

注2　将一个 g_0-随机数 v 在取舍原则中"被选用"的概率记为 p,将 g_0-随机数列经取舍原则而首次接受时,已取舍的次数记为 N.那么

$$p = P\left(V < +\infty, \frac{f(V)}{Cg_0(V)} \geqslant U\right) = \int P\left(\frac{f(V)}{Cg_0(V)} \geqslant U \mid V = v\right) g_0(v)\mathrm{d}v$$
$$= \int \frac{f(v)}{Cg_0(v)} g_0(v)\mathrm{d}v = \frac{1}{C}.$$

而且 N 服从参数为 p 的几何分布,因此得到下面的命题.

命题1.4　在利用 $f(x) \leqslant Cg_0(x)$ 作为取舍原则的根据时,为了得到一个 f-随机数,平均需要 C 个 g_0-随机数.

证明　$EN = \dfrac{1}{p} = C.$

由此可见 C 取得越小,取舍的效率越高.

1.2.3　对于取舍原则的再认识

下面将 $g_0(x)$ 简记为 $g(x)$.因为 $g(x)$ 是一个密度函数,又假定了 Y,U 相互独立,$U \sim U[0,1]$,$Y \sim g(x)$.取舍原则是说,一个 Y-随机数 y 被接受,当且仅当对于 U-随机数 u 有 $f(y) \geqslant Cg(y)u$.我们再定义一个新的随机变量

$$Z = Cg(Y)U.$$

于是,在 $Y=y$ 条件下,Z 的条件密度是均匀分布 $f_{Z|Y}(z|y) = \dfrac{1}{Cg(y)} I_{[0,Cg(y)]}(z)$.故而 (Y,Z) 的联合密度是

$$f_{Y,Z}(y,z) = g(y) \cdot \frac{1}{Cg(y)} I_{[0,Cg(y)]}(z) = \frac{1}{C} I_{[0,Cg(y)]}(z).$$

即二元随机向量 (Y,Z) 服从集合 $\{(y,z): 0 \leqslant z \leqslant Cg_0(y)\}$ 上均匀分布.从而取舍原则可以重述为

一个 Y-随机数 y 被接受,当且仅当 $f(y) \geqslant z$　(z 是一个 Z-随机数).

1.2.4　取舍原则的另一种表述

找密度函数 $f(x)$ 的一个控制函数 $M(x) \geqslant f(x)$,使二维集合 $\{(y,z): 0 \leqslant z \leqslant M(y)\}$ 上的均匀随机数容易生成(例如,经典情形就是 $M(x) = Cg(x)$).于是取舍原则可以叙述为:

生成 (Y,Z)-随机数 (y,z),如果 $f(y) \geqslant z$,则保留 y,否则舍弃 y.那么这样保留下的 y,就是 f-随机数.

证明　被接受的 Y-随机数的分布函数是

$$P(Y \leqslant x \mid (Y,Z): 0 \leqslant Z \leqslant f(Y)) = \frac{P((Y,Z): Y \leqslant x, 0 \leqslant Z \leqslant f(Y))}{P((Y,Z): 0 \leqslant Z \leqslant f(Y))}$$

$$= \frac{\int_{-\infty}^{x}\int_{0}^{f(y)} \mathrm{d}z\mathrm{d}y}{\int_{-\infty}^{+\infty}\int_{0}^{f(y)} \mathrm{d}z\mathrm{d}y} = \frac{\int_{-\infty}^{x} f(y)\mathrm{d}y}{\int_{-\infty}^{+\infty} f(y)\mathrm{d}y} = \int_{-\infty}^{x} f(y)\mathrm{d}y.$$

1.3　Gauss 系

1.3.1　方差有限的随机变量全体组成的 Hilbert 空间

将概率空间 $(\Omega, \mathscr{F}, P)$ 上有方差的全体随机变量的集合（依赖于事件体 $\mathscr{F}$），记为 $L^2(\Omega)$. 显见它是 Euclid 空间，即它是线性空间，并且可以定义**内积**：

$$(\xi, \eta) \stackrel{\text{def}}{=\!=} E(\xi\eta).$$

$L^2(\Omega)$ 作为线性空间，与线性代数中大多数 Euclid 空间的例子有本质的不同，就在于它不是有限维的，即在 $L^2(\Omega)$ 中找不到有限个元，使之构成一组基. 实际上，在 $L^2(\Omega)$ 中可以找到无穷多个线性无关的元素，这样的 Euclid 空间，称为**无穷维 Euclid 空间**.

对于 $L^2(\Omega)$ 中任意元 ξ，可以定义其**范数**为

$$\| \xi \| \stackrel{\text{def}}{=\!=} \sqrt{E\xi^2}.$$

而对于任意 $\xi, \eta \in L^2(\Omega)$，差的范数 $\| \xi - \eta \|$，称为 ξ, η 间的**距离**. 可以证明，这个距离和普通的三维空间中两点之间的距离非常相似，即它们满足三角形不等式

$$\| \xi - \zeta \| \leqslant \| \xi - \eta \| + \| \eta - \zeta \|.$$

在 $L^2(\Omega)$ 中的随机变量序列 $\{\xi_n\}$，称为收敛到 $\xi \in L^2(\Omega)$，如果当 $n \to +\infty$ 时有

$$\| \xi_n - \xi \| \to 0.$$

它等价于

$$E(\xi_n - \xi)^2 \to 0.$$

一般将此种收敛，称为随机变量列的**均方收敛**（参见 1.1.4 节）. 概率理论指出，这种收敛与概率为 1 地收敛，即

$$P(\omega: \xi_n(\omega) \to \xi(\omega)) = 1,$$

互不蕴含. 但是，它们都蕴含依概率收敛.

定义 1.3　在 $L^2(\Omega)$ 中的随机变量序列 $\{\xi_n\}$，称为 Cauchy 列，如果当 $n, m \to +\infty$ 时，恒有

$$\| \xi_n - \xi_m \| \to 0.$$

我们不加证明地引用下述基本结论

命题 1.5　$L^2(\Omega)$中的任意 Cauchy 列必有极限，即如果 $L^2(\Omega)$中的随机变量序列$\{\xi_n\}$，满足：当 $n,m\to+\infty$时恒有 $\|\xi_n-\xi_m\|\to 0$，那么，一定在 $L^2(\Omega)$中存在 ξ，使$\|\xi_n-\xi\|\to 0$.

定义 1.4　如果无穷维 Euclid 空间中的任意 Cauchy 列都有极限，则称为此 Euclid 空间为**完备的**. 完备的 Euclid 空间通常称为 **Hilbert 空间**.

所以，$L^2(\Omega)$是 Hilbert 空间.

用 Hilbert 空间的观点理解随机变量的均方收敛性，有与三维空间类似的直观感觉.

1.3.2　作为 d 维正态分布的推广的 d 维 Gauss 分布

定义 1.5　如果存在 m 个独立的标准正态随机变量 $\eta_1,\eta_2,\cdots,\eta_m$，以及常数 a_{ij} $(1\leqslant i\leqslant d, 1\leqslant j\leqslant m)$与 μ_i $(1\leqslant i\leqslant d)$，使得

$$\xi_1 = a_{11}\eta_1+\cdots+a_{1m}\eta_m+\mu_1$$
$$\vdots$$
$$\xi_d = a_{d1}\eta_1+\cdots+a_{dm}\eta_m+\mu_d,$$

则称 d 维随机向量$\boldsymbol{\xi}=(\xi_1,\xi_2,\cdots,\xi_d)^{\mathrm{T}}$(上角 T 表示转置运算)服从 **$d$ 维 Gauss 分布**.

矩阵形式为

$$\boldsymbol{\xi}=\boldsymbol{A}\boldsymbol{\eta}+\boldsymbol{\mu},$$

其中

$$\boldsymbol{A}=(a_{ij})_{d\times m}.$$

这时期望向量与协方差矩阵分别为

$$E\boldsymbol{\xi}=\boldsymbol{\mu},$$
$$\boldsymbol{\Sigma}=E(\boldsymbol{\xi}-\boldsymbol{\mu})(\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}.$$

协方差矩阵$\boldsymbol{\Sigma}$为可逆时的 Gauss 分布，称为**(多维)正态分布**，此时存在分布密度函数

$$f(\boldsymbol{x})=\frac{1}{(2\pi)^{\frac{d}{2}}\,|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}.$$

$\boldsymbol{\xi}$服从 d 维 Gauss 分布，记为$\boldsymbol{\xi}\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$.

Gauss 分布是正态分布的自然推广. 例如，常数 a 看成随机变量并不服从正态分布，因为它的方差是 0. 但是，按定义它是一维 Gauss 分布.

正态分布就是不退化的 Gauss 分布，只当此时才存在分布密度. 一般的 Gauss 分布可能只分布在一个低维的超平面上.

服从 Gauss 分布的随机向量，称为 **Gauss 随机向量**. 由它的定义直接可知，Gauss 随机向量的任意多个分量仍然是 Gauss 随机向量，或 Gauss 随机变量.

Gauss 分布有下列基本事实

(1) 若服从 d 维 Gauss 分布$\boldsymbol{\xi}\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$，则此时，对于 $m\times d$ 矩阵$\boldsymbol{C}$，及 m 维列向量

$\boldsymbol{b}$ 有

$$\boldsymbol{C\xi}+\boldsymbol{b}\sim N(\boldsymbol{C\mu}+\boldsymbol{b},\boldsymbol{C\Sigma C}^{\mathrm{T}}).$$

(2) 若$\boldsymbol{\xi}_1\sim N(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)$,$\boldsymbol{\xi}_2\sim N(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$,且相互独立,则

$$\boldsymbol{\xi}_1+\boldsymbol{\xi}_2\sim N(\boldsymbol{\mu}_1+\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2).$$

(3) 一维 Gauss 随机变量 ξ 或者是正态随机变量,或者是常数. 此时有

$$Ee^{\xi}=e^{E\xi+\frac{1}{2}\mathrm{var}(\xi)}.$$

(4) $\{\boldsymbol{\xi}^{(n)}\}$服从 Gauss 分布,且$\boldsymbol{\xi}^{(n)}\xrightarrow{p}\boldsymbol{\xi}$(指所有分量都依概率收敛),则$\boldsymbol{\xi}$服从 Gauss 分布.

(5) Gauss 随机向量对依分布收敛的封闭性. 即若$\boldsymbol{\xi}_n\sim N(\boldsymbol{\mu}_n,\boldsymbol{\Sigma}_n)$,且$\boldsymbol{\xi}_n\xrightarrow{d}\boldsymbol{\xi}$ (指其相应的特征函数$\varphi_{\boldsymbol{\xi}_n}(\boldsymbol{\lambda})\stackrel{\mathrm{def}}{=\!=}Ee^{i\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_n}\to\varphi_{\boldsymbol{\xi}}(\boldsymbol{\lambda})\stackrel{\mathrm{def}}{=\!=}Ee^{i\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}}$),则存在$\boldsymbol{\mu}$,$\boldsymbol{\Sigma}$,使

$$\boldsymbol{\mu}_n\to\boldsymbol{\mu},\boldsymbol{\Sigma}_n\to\boldsymbol{\Sigma},\quad 而且\quad \boldsymbol{\xi}\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma}).$$

命题 1.6 下面 4 个叙述彼此等价:

(1) 随机向量$\boldsymbol{\xi}=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_d\end{pmatrix}$服从 d 维 Gauss 分布.

(2) 对于任意实数 $a_1,a_2,\cdots,a_d$,线性组合$\sum\limits_{k=1}^{d}a_k\xi_k$ 服从一维 Gauss 分布.

(3) $\boldsymbol{\xi}$ 的特征函数$\varphi(\boldsymbol{\lambda})\stackrel{\mathrm{def}}{=\!=}Ee^{i\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}}$有以下形式

$$\varphi(\boldsymbol{\lambda})=e^{i\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\mu}-\frac{1}{2}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Sigma\lambda}},$$

(4) $\boldsymbol{\xi}$ 的矩母函数$M(\boldsymbol{z})\stackrel{\mathrm{def}}{=\!=}Ee^{\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\xi}}$有以下形式

$$M(\boldsymbol{z})=e^{\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\Sigma z}}.$$

1.3.3 Gauss 系与 Gauss 过程

定义 1.6 随机变量族$\{\xi_t: t\in I\}$称为 **Gauss 系**,如果对于任意 n 及任意 $t_1,t_2,\cdots,t_n\in I$,随机向量$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$服从 Gauss 分布.

对于 Gauss 系$\{\boldsymbol{\xi}^{(n)}\}$而言,依概率收敛与均方收敛是等价的.

如果 Gauss 系中的指标集 $I=[0,+\infty)$,则称为 **Gauss 过程**.

定义 1.7 随机过程$\{\xi_t: t\geqslant 0\}$对于每一个固定的基本事件(样本点)ω 作为 t 的函数 $\xi_t(\omega)$,称为随机过程$\{\xi_t: t\geqslant 0\}$对应于样本点 ω 的**样本轨道**. 如果随机过程$\{\xi_t: t\geqslant 0\}$具有有限方差,即对于任意 $t\geqslant 0$ 恒有 $\mathrm{var}(\xi_t)<+\infty$. 记

$$\mu(t)=E\xi_t,\quad R(s,t)=E(\xi_s\xi_t),$$

$$C(s,t) = \operatorname{cov}(\xi_s,\xi_t) = R(s,t) - \mu(s)\mu(t).$$

它们分别称为随机过程$\{\xi_t: t\geqslant 0\}$的**均值函数**、**相关函数**与**协方差函数**.

显见

$$C(s,t) = R(s,t) - \mu(s)\mu(t).$$

定理 1.7 随机过程$\{\xi_t: t\geqslant 0\}$是 Gauss 过程,当且仅当它在任意有限个时刻的任意线性组合的分布都是 Gauss 分布. 即对于任意 n,任意实数 $a_1,a_2,\cdots,a_n$ 及任意时刻 $t_1,t_2,\cdots,t_n$,随机变量 $a_1\xi_{t_1}+a_2\xi_{t_2}+\cdots+a_n\xi_{t_n}$ 或者都是常数,或者服从正态分布.

对于 Gauss 过程,其均值函数与相关函数完全地确定了它的**有限维分布族**,即对于任意 n 及任意 $t_1,t_2,\cdots,t_n\in I$,随机向量$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$的联合分布全体:

$$\begin{pmatrix}\xi_{t_1}\\ \xi_{t_2}\\ \vdots\\ \xi_{t_n}\end{pmatrix} \sim N\left(\begin{pmatrix}\mu(t_1)\\ \mu(t_2)\\ \vdots\\ \mu(t_n)\end{pmatrix}, \begin{pmatrix} \boldsymbol{C}(t_i,t_j) \end{pmatrix}_{i,j\leqslant n}\right),$$

其中 $C(s,t)$是协方差函数. 可见,Gauss 过程的概率特性完全由其均值函数和协方差函数所决定.

定义 1.8 如果随机系$\{\xi_t: t\in I\}$具有有限方差,那么,$L^2(\Omega)$中的如下子集 $L(\xi)\overset{\text{def}}{=\!=}\{\xi_t: t\in I\}$中任意有限个元素的实线性组合组成的集合;称为$\{\xi_t: t\in I\}$的**线性包**. 再记 $\bar{L}(\xi)$为包含 $L(\xi)$,且对于均方极限封闭的最小集合:即

$$\text{若 } \eta^{(n)} \in \bar{L}(\xi), E\mid \eta^{(n)} - \eta\mid^2 \to 0, \text{则 } \eta\in \bar{L}(\xi).$$

那么,$\bar{L}(\xi)$是$\{\xi_t: t\in I\}$的**线性闭包**,称为随机系$\{\xi_t: t\in I\}$的**线性均方信息空间**. 线性均方信息空间中的每个元素都是$\{\xi_t: t\in I\}$中元素线性组合在均方意义下的极限,因而可看成整个随机变量族$\{\xi_t: t\in I\}$的某个“线性泛函”.

命题 1.8(封闭性) 如果$\{\xi_t: t\in I\}$是 Gauss 系,则其均方线性空间 $\bar{L}(\xi)$也是 Gauss 系.

此外,我们还有下面的命题.

命题 1.9(独立性)

(1) 若$\{\xi_\alpha: \alpha\in I\}$与$\{\eta_\beta: \beta\in J\}$独立,那么 $\bar{L}(\xi)$与 $\bar{L}(\eta)$独立(指任意各自有限个元素组成的随机向量都相互独立).

(2) 若$\{\xi_\alpha,\eta_\beta: \alpha\in I,\beta\in J\}$是 Gauss 系,则$\{\xi_\alpha: \alpha\in I\}$与$\{\eta_\beta: \beta\in I\}$独立的充要条件为对于任意 ξ_α,η_β 都有 $\operatorname{cov}(\xi_\alpha,\eta_\beta)=0$.

(3) 若 Gauss 系$\{\xi_\alpha: \alpha\in I\}$与 Gauss 系$\{\eta_\beta: \beta\in J\}$相互独立,那么$\{\xi_\alpha,\eta_\beta: \alpha\in I,\beta\in J\}$也是 Gauss 系.

定理 1.10 设$\{\xi_t^{(n)}: t\geqslant 0\}$是一系列 Gauss 过程. 又对于任意 t 都有

$$E \mid \xi_t^{(n)} - \xi_t \mid^2 \xrightarrow{n \to +\infty} 0$$

那么，$\{\xi_t: t \geqslant 0\}$也是 Gauss 过程.

证明　可以直接验证，对于任意 m 及任意时刻 $t_1, t_2, \cdots, t_m$，$(\xi_{t_1}^{(n)}, \xi_{t_2}^{(n)}, \cdots, \xi_{t_m}^{(n)})$的联合特征函数趋于$(\xi_{t_1}, \xi_{t_2}, \cdots, \xi_{t_m})$的联合特征函数. 由假定$(\xi_{t_1}^{(n)}, \xi_{t_2}^{(n)}, \cdots, \xi_{t_m}^{(n)})$是 Gauss 分布，因此$(\xi_{t_1}, \xi_{t_2}, \cdots, \xi_{t_m})$是 Gauss 分布. 故而$\{\xi_t: t \geqslant 0\}$是 Gauss 过程. .

复 Gauss 过程

设 $\zeta_t^{(k)} = \xi_t^{(k)} + \mathrm{i}\eta_t^{(k)}$ $(k=1,2)$中实部与虚部的期望为 0，则二元复函数

$$R(s,t) \stackrel{\text{def}}{=} E(\zeta_s^{(1)} \overline{\zeta_t^{(2)}})$$

称为复随机过程$\{\zeta_t^{(1)}: t \geqslant 0\}$与$\{\zeta_t^{(2)}: t \geqslant 0\}$的**相关函数**. 它是一个复值的非负定函数，即对于任意 $m, t_1, t_2, \cdots, t_m$ 及任意复数 $\alpha_1, \alpha_2, \cdots, \alpha_m$，恒有

$$\sum_{k,l=1}^{m} R(t_k, t_l) \alpha_k \overline{\alpha_l} \geqslant 0.$$

定义 1.9　复值的随机过程$\{\zeta_t = \xi_t + \mathrm{i}\eta_t: t \geqslant 0\}$，称为**复 Gauss 过程**，如果实值过程$\{\xi_t: t \geqslant 0\}$与$\{\eta_t: t \geqslant 0\}$是相互独立的 Gauss 过程，而且有相同的有限维分布族.

习题 1

1. 证明：(1) $|P(AB) - P(A)P(B)| \leqslant \frac{1}{4}$；

(2) $|P(AB) - P(BC)| \leqslant 1 - P(AC)$；

(3) $P(A_1 \cdots A_n) \geqslant 1 - \sum_{k=1}^{n} (1 - P(A_k))$.

2. 由 $P(A|B) > P(A)$推出 $P(B|A) > P(B)$，解释其含义.

3. 设 X 为取非负整数值的随机变量，证明：

(1) $EX = \sum_{n=1}^{+\infty} P(X \geqslant n)$；

(2) $\mathrm{var}(X) = 2\sum_{n=1}^{+\infty} nP(X \geqslant n) - EX(EX+1)$.

4. 设随机变量 X 与 Y 独立，且方差有限，证明：

(1) $\mathrm{var}(XY) = \mathrm{var}(X)\mathrm{var}(Y) + (EX)^2\mathrm{var}(Y) + (EY)^2\mathrm{var}(X)$；

(2) $\mathrm{var}(XY) \geqslant \mathrm{var}(X)\mathrm{var}(Y)$.

5. 设 $\xi_1, \xi_2, \cdots, \xi_{n+m}$ $(n > m)$ 相互独立同分布且具有有限方差，试求 $X = \sum_{k=1}^{n} \xi_k$ 与 $Y = \sum_{k=1}^{n} \xi_{m+k}$ 的相关系数.

6. 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，求 $E|X-\mu|$.

7. 设随机变量 X 服从参数为 λ 的 Poisson 分布，求 $E|X-\mu|$.

8. 设随机变量 X 的概率密度函数为 $f(x)$，令 $h(v)=E|X-v|$，试证明：当 v 满足 $P(X\leqslant v)=\frac{1}{2}$ 时(v 称为 X 的中位数)，$h(v)$ 达到最小.

9. 零件强度 $X\sim N(48,2^2)$. 若 $X\geqslant 46.32$ 为合格品，否则为次品. 今检验方案为：第一次任取3个，若3个都合格，则接收该批零件；若至少有2个次品，则拒收该批零件；否则再抽第二次，也是3个. 重复上述步骤，至作出接收或拒绝决定为止.

(1) 求由第一次取样而接收该批零件的概率；

(2) 求该批产品被接收的概率；

(3) 记 N 为在作出决定时的抽取次数，求 N 的分布.

10. 设 $\tau\sim \mathrm{Exp}_\lambda$，求 $\tau\wedge t$ 的分布函数.

11. 设 $\xi\sim \mathrm{Exp}_\lambda$，$\eta\sim \mathrm{Exp}_\lambda$，$\xi,\eta$ 独立，求 $P(\xi>\eta)$ 并求 $\xi\wedge\eta$ 及 $(\xi-\eta)^+$ 的分布.

12. 设 $F(x)$ 为分布函数，$\bar{F}(x)=1-F(x)$. 证明 $F(x)$ 是指数分布的充要条件为

$$\bar{F}(t+s)=\bar{F}(s)\bar{F}(t).$$

13. 设随机变量 X 服从 $N(0,1)$ 分布. 求 X 与 X^n(n 为正整数)的协方差与相关系数.

14. 假设 $E(X|Y)=EX$，证明 X 与 Y 不相关. 举例说明其逆命题不成立.

15. 设 $X_1,X_2,\cdots,X_n$ 为 n 个具有密度 $f(x)$(分布函数记为 $F(x)$)的独立同分布随机变量. 令 $X_{(1)}$ 为它们中最小的值，即 $X_{(1)}=\min\limits_{k\leqslant n}X_k$，$X_{(2)}$ 为它们中第二个小的值，……，$X_{(n)}$ 为它们中最大的值.

(1) 证明 $X_{(k)}$ 具有密度 $kC_n^k f(x)F(x)^{k-1}(1-F(x))^{n-k}$.

(2) 若独立同分布的 $X_1,X_2,\cdots,X_n\sim U[0,1]$，求 $EX_{(k)}$，$\mathrm{var}(X_{(k)})$ 及相关系数 $\rho(X_{(1)},X_{(n)})$. 又问 $\mathrm{var}(X_{(1)}),\mathrm{var}(X_{(2)}),\cdots,\mathrm{var}(X_{(n)})$ 中哪个最小，哪个最大？证明当 $n\to+\infty$ 时，$X_{(n)}$ 依概率收敛到1，且 $n(1-X_{(n)})$ 依分布收敛到 Exp_1.

16. (1) 证明 $\mathrm{var}(Y)\geqslant E[\mathrm{var}(Y|X)]$，其中 $\mathrm{var}(Y|X)=E[(Y-E(Y|X))^2|Y]$ 称为 Y 对 X 的条件方差.

(2) 在 (X,Y) 服从二维正态分布时，求证 $\mathrm{var}(Y|X)=(1-\rho^2)\mathrm{var}(Y)$，其中 ρ 是 X 与 Y 的相关系数.

17. 若系统由两种独立工作的设备串联而成，它们的寿命分别为参数 λ,μ 的指数分布，又它们各有备件 $n-1$ 件与 $m-1$ 件. 求在这些备件的支持下，系统的平均寿命.

18. 设 A,B 为随机事件. 已知 $P(A),P(B),P(AB)$，求 $E(I_B|A)$，$E(I_{A\cup B}|A)$ 及 $E(I_{A\cap B}|A)$，并分别求它们的分布.

19. 若 $\psi(t)\overset{\mathrm{def}}{=\!=}E\mathrm{e}^{tX}<+\infty$，证明以下的 Chernoff 不等式：$\forall t$

$$P(X>a)\leqslant \mathrm{e}^{-(at-\psi(t))}.$$

再由此推出

$$P(X>a)\leqslant e^{-\sup_t(at-\psi(t))}.$$

20. 若 $g(x)$,$h(x)$都是非增函数,或都是非降函数.证明 $E[g(X)h(X)]\geqslant Eg(X)Eh(X)$,并解释其含义.又若它们一个非增,另一个非降,则结果如何?

21. 设随机变量 $\tau_1,\tau_2,\cdots,\tau_n$ 相互独立,且 $\tau_k\sim \mathrm{Exp}_{\lambda_k}(k\leqslant n)$.证明

$$P(\tau_j=\max_{k\leqslant n}\tau_k)=\frac{\lambda_j}{\lambda_1+\lambda_2+\cdots+\lambda_n}.$$

22. 设随机变量 $X\sim U[-\pi,\pi]$,$S_n=\sum_{k=1}^{n}\cos(kX)$.证明$\frac{S_n}{n}\xrightarrow{p}0(n\to+\infty)$.

23. 设 X,Y,Z,W 是独立同分布的标准正态随机变量.证明

$$P(a\sqrt{X^2+Y^2}>b\sqrt{Z^2+W^2})=\frac{b^2}{a^2+b^2}.$$

24. 设 $F(x)$,$G(y)$是两个分布函数.记

$$F_1(x,y)=(F(x)+G(y)-1)^+,\quad F_2(x,y)=\min\{F(x),G(y)\},$$

其中记号 $z^+=zI_{[0,+\infty)}(z)$.证明 $F_1(x,y)$,$F_2(x,y)$都是二元分布函数,且满足

(1) $F_1(x,+\infty)=F_2(x,+\infty)=F(x)$,$F_1(+\infty,y)=F_2(+\infty,y)=G(y)$;

(2) 若另有二元分布函数 $F(x,y)$,也满足 $F(x,+\infty)=F(x)$,$F(+\infty,y)=G(y)$,则 $F_1(x,y)\leqslant F(x,y)\leqslant F_2(x,y)$.

25. 移位指数分布的密度为 $p(x)=\lambda e^{-\lambda(x-a)}I_{[a,+\infty)}(x)$,如何得到它的样本?

26. 若随机变量 $\xi\sim N(0,1)$,$\eta\sim\chi^2(n)$且与 ξ 独立,则 $\dfrac{\xi}{\sqrt{\frac{\eta}{n}}}$服从 $t(n)$分布 $\left(\text{其密度为 } t(n,x)=\dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\dfrac{x^2}{n}\right)^{-\frac{n+1}{2}}\right)$.如何构造 $t(n)$随机数?

27. 求密度为$\dfrac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}a^{-\frac{1}{2}}\left(1+\dfrac{x^2}{ak}\right)^{-\frac{k+1}{2}}$的 $t(k,a)$分布的方差.如何取它的随机数?

28. 设 ξ_t 是轨道为 t 的连续函数的 Gauss 过程,$E\xi_t=m(t)$,$\mathrm{cov}(\xi_s,\xi_t)=R(s,t)-m(s)m(t)$,$G(t)$是单调递增函数.如果存在随机变量 η_t,使 $\mathrm{var}(\eta_t)<+\infty$,而且当 $\max_i\{t_{i+1}^{(n)}-t_i^{(n)}\}\xrightarrow{n\to+\infty}0$ 时,有

$$E\left|\sum_i\xi_{t_i^{(n)}}[G(t_{i+1}^{(n)})-G(t_i^{(n)})]-\eta_t\right|^2\to 0,$$

这里$\{t_i^{(n)}\}$是半开区间$(0,t]$的一个划分. 那么,定义$\int_0^t \xi_s \mathrm{d}G(s) = \eta_t$. 证明$\eta_t$也是 Gauss 过程. 再求$E\eta_t$,$\mathrm{cov}(\eta_s,\eta_t)$.

29. 设ξ_t,η_t是相互独立的 Gauss 过程,证明$\xi_t+\eta_t$也是 Gauss 过程.

30. 设ξ_t是轨道为t的连续函数的 Gauss 过程,$E\xi_t=m(t)$,$\mathrm{cov}(\xi_s,\xi_t)=R(s,t)$,$G(t)$是任意单调递增函数. Gauss 过程$\xi_t$的特征泛函定义为

$$\Phi_\xi(G) = E\mathrm{e}^{\mathrm{i}\int_0^t \xi_s \mathrm{d}G(s)}.$$

从这个特征泛函的显式表示,对于$t_1<t_2<t$,取$G(t)=\theta_1 I_{[0,t_1]}(t)+\theta_2 I_{(0,t_2]}(t)$,证明得到的$\Phi_\xi(G)$正是$(\xi_{t_1},\xi_{t_2})$的联合特征函数$\varphi_{\xi_{t_1},\xi_{t_2}}(\theta_1,\theta_2)=E\mathrm{e}^{\mathrm{i}(\theta_1\xi_{t_1}+\theta_2\xi_{t_2})}$的显式表示.

第 2 章　条件分布与条件期望

2.1　条件分布与全概率公式的推广

2.1.1　条件分布

设随机向量(ξ,η)能取的值为$(x_i,y_j)(i,j=1,2,\cdots)$. 那么,在$\eta$取值$y$的条件下,$\xi$的条件分布是指下面的分布表

$$\begin{pmatrix} \cdots & x_i & \cdots \\ \cdots & P(\xi=x_i \mid \eta=y) & \cdots \end{pmatrix},$$

其中

$$P(\xi=x_i \mid \eta=y)=\begin{cases} P(\xi=x_i \mid \eta=y_j), & y=y_j, \quad j=1,2,\cdots, \\ 0, & y\neq y_j, \quad j=1,2,\cdots \end{cases}$$

是条件概率.

设随机向量$(\xi,\eta)\sim$联合密度$f(x,y)$,则在$\eta=y$的条件下,随机变量ξ的条件密度,记为$f_\xi(x|\eta=y)$,或$f_{\xi|\eta}(x|y)$,其表示式为

$$f_{\xi|\eta}(x \mid y)=f_\xi(x \mid \eta=y)=\frac{f(x,y)}{f_\eta(y)},$$

其中$f_\eta(y)=\int f(x,y)\mathrm{d}x$是$\eta$的(边缘)分布密度.

2.1.2　全概率公式及其积分形式

全概率公式　若正概率事件序列$\{A_n\}_{n\geqslant 1}$满足

$$\forall A_n\in\mathscr{F} \quad (n\geqslant 1), \quad \bigcup_{n\geqslant 1}A_n=\Omega, \quad A_n\cap A_m=\varnothing \quad (n\neq m),$$

那么,对任意$B\in\mathscr{F}$,有

$$P(B)=\sum_{n\geqslant 1}P(B \mid A_n)P(A_n).$$

我们可以直接从定义验证下面的公式.

积分形式的全概率公式　设随机向量$(\xi,\eta)\sim$联合密度$f(x,y)$,则对于任意随机事件A有

$$P(A)=\int P(A\mid\eta=y)f_\eta(y)\mathrm{d}y.$$

特别地，对于 Borel 事件体的任意集合(Borel 集)Λ 有

$$P(\xi^{-1}\Lambda)\equiv P(\xi\in\Lambda)=\int P(\xi\in\Lambda\mid\eta=y)f_\eta(y)\mathrm{d}y.$$

它们都是后面将论述的全期望公式的特殊情形.

2.1.3　Bayes 公式及其积分形式

Bayes 公式

$$P(A_i\mid B)=\frac{P(B\mid A_i)P(A_i)}{P(B)}.$$

积分形式的 Bayes 公式

设随机向量$(\xi,\eta)\sim$联合密度 $f(x,y)$，则

$$f_\eta(y\mid\xi=x)=\frac{f_\xi(x\mid\eta=y)f_\eta(y)}{\int f_\xi(x\mid\eta=z)f_\eta(z)\mathrm{d}z}.$$

2.1.4　多维条件分布

设随机向量$(\boldsymbol{\xi},\boldsymbol{\eta})\sim$联合密度 $f(\boldsymbol{x},\boldsymbol{y})$，那么

$$f_{\boldsymbol{\xi}}(\boldsymbol{x}\mid\boldsymbol{y})\stackrel{\text{def}}{=}f_{\boldsymbol{\xi}}(\boldsymbol{x}\mid\boldsymbol{\eta}=\boldsymbol{y})=\frac{f(\boldsymbol{x},\boldsymbol{y})}{f_{\boldsymbol{\eta}}(\boldsymbol{y})},$$

其中 $f_{\boldsymbol{\eta}}(\boldsymbol{y})=\int f(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{x}$ 是$\boldsymbol{\eta}$ 的(边缘)分布.

例 2.1　$\begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}\sim N\left(\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix},\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}\right)$，且$\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}$与$\boldsymbol{\Sigma}_{22}$都可逆，则直接计算可以得到条件分布 $f_{\boldsymbol{\xi}}(\boldsymbol{x}|\boldsymbol{y})$是正态分布 $N(\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_2),\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$.

2.2　条件期望

2.2.1　随机变量在另一个随机变量取定值时的数学期望

在 $\eta=y$ 条件下，ξ 的条件分布的期望为

(1) 若(ξ,η)为离散型的随机向量，则

$$E(\xi\mid\eta=y)=\sum_x xP(\xi=x\mid\eta=y)=\frac{\sum_x xf(x,y)}{\sum_x f(x,y)}.$$

如果 η 可能取的值为 $y_1, y_2, \cdots, y_j, \cdots$，那么，上式又可以写为

$$E(\xi \mid \eta = y) = \sum_j \frac{\sum_x x f(x, y)}{\sum_x f(x, y)} I_{\{y_j\}}(y).$$

(2) 若(ξ, η)为连续型的随机向量，则

$$E(\xi \mid \eta = y) = \int x f_\xi(x \mid y) \mathrm{d}x = \frac{\int x f(x, y) \mathrm{d}x}{\int f(x, y) \mathrm{d}x}.$$

可见，在条件 $\eta = y$ 下的条件期望 $E(\xi|\eta = y)$是 y 的一个函数，我们将它记为 $\varphi(y)$.

2.2.2　随机变量对另一个随机变量的条件期望

我们用随机变量 $\varphi(\eta)$来统一表达在 η 取不定值(随机值)时，ξ 的条件期望，并将它记为 $E(\xi|\eta)$，即 $E(\xi|\eta) \overset{\text{def}}{=\!=} \varphi(\eta)$，其中 $\varphi(y) = E(\xi|\eta = y)$.

于是 $E(\xi|\eta)$是一个随机变量，它是 η 的函数，即

当 $\eta = y$ 时，它的值为 $E(\xi \mid \eta = y)$.

所以，$E(\xi|\eta)$称为 ξ 关于 η 的**条件期望**.

例 2.2　设$\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}\right)$，而且$\begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$与$\boldsymbol{\Sigma}_{22}$都可逆，则 $E(\boldsymbol{\xi}|\boldsymbol{\eta} = \boldsymbol{y})$为条件分布 $f_{\boldsymbol{\xi}}(\boldsymbol{x}|\boldsymbol{y})$的期望，也就是正态分布 $N(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$的期望. 故而

$$E(\boldsymbol{\xi} \mid \boldsymbol{\eta} = \boldsymbol{y}) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}_2).$$

于是

$$E(\boldsymbol{\xi} \mid \boldsymbol{\eta}) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} (\boldsymbol{\eta} - \boldsymbol{\mu}_2),$$

它是$\boldsymbol{\eta}$的线性函数.

例 2.3　设(ξ, η)为离散型随机向量，则

$$E(\xi \mid \eta) = \frac{\sum_x x f(x, \eta)}{\sum_x f(x, \eta)} = \sum_j \frac{\sum_x x f(x, \eta)}{\sum_x f(x, \eta)} I_{\{y_j\}}(\eta)$$

例 2.4　设(ξ, η)为连续型随机向量，则

$$E(\xi \mid \eta) = \frac{\int x f(x, \eta) \mathrm{d}x}{\int f(x, \eta) \mathrm{d}x}.$$

(1) 条件期望的含义

随机变量 ξ 关于随机变量 η 的条件期望 $E(\xi|\eta)$满足以下的基本性质：

(E.1) 随机变量 ξ 关于随机变量 η 的条件期望 $E(\xi|\eta)$ 是 η 的函数.

(E.2) 对任意 Borel 集 Λ，有

$$E[I_\Lambda(\eta)E(\xi \mid \eta)] = E(\xi \cdot I_\Lambda(\eta)),$$

其中

$$I_\Lambda(x) = \begin{cases} 1, & x \in \Lambda, \\ 0, & x \notin \Lambda \end{cases}$$

是集合 Λ 的示性函数.

我们在 $(\xi,\eta) \sim$ 联合密度 $f(x,y)$ 情形，验证性质(E.2)如下：(E.2)的左边是

$$\int I_\Lambda(y)E(\xi \mid \eta = y)f_\eta(y)\mathrm{d}y = \int I_\Lambda(y)\frac{\int xf(x,y)\mathrm{d}x}{\int f(x,y)\mathrm{d}x}\left(\int f(x,y)\mathrm{d}x\right)\mathrm{d}y$$

$$= \int I_\Lambda(y)xf(x,y)\mathrm{d}x\mathrm{d}y = \text{(E.2) 的右式.}$$

此外，概率理论证明了，由(E.1)及(E.2)确定的随机变量，在不计零概率事件的差别下，是唯一的.

注　上面叙述的是在 (ξ,η) 为连续型的或离散型的情形时的性质. 事实上，对于一般的任意随机变量 (ξ,η)，概率理论证明了：只要随机变量 ξ 的数学期望存在，那么就一定存在随机变量 η 的某个函数 $\psi(\eta)$，使得对实数 Borel 集 Λ，恒有

$$E(I_\Lambda(\eta)\psi(\eta)) = E(\xi \cdot I_\Lambda(\eta)).$$

这时，就将 $\psi(\eta)$ 定义为随机变量 ξ 关于随机变量 η 的条件期望，并仍旧记之为 $E(\xi|\eta)$. 然而这样的定义是较为抽象的，一般很难得到函数 ψ 的表达形式.

(E.1)与(E.2)的含义为：$E(\xi|\eta)$ 是这样一个随机的量，它在 η 取值已知时是完全确定的，称为 η **可知的**，而且在有关 η 的任意一个随机事件 $\{\eta \in \Lambda\}$ 上看，它的平均与 ξ 的平均是相等的.

条件期望 $E(\xi|\eta)$ 是一个随机变量，其直观含义为：将随机变量 η 看成固定时，在顾及随机变量 ξ 与随机变量 η 间的联系时 ξ 的数学期望.

显见，对于常数 c 有

$$E(c \mid \eta) = c.$$

(2) 条件期望的运算规律

条件期望的重要运算规律如下.

定理 2.1(全期望公式)

$$E(E(\xi \mid \eta)) = E\xi.$$

(只需在(E.2)中将集合 Λ 取为实数全集，就得到这个公式).

在应用中的计算形式，实际上是将 $E\xi$ 表达为 $E(E(\xi|\eta))$. 最常见情形是，随机变量 η

具有密度函数 $f_\eta(y)$ 的情形，此时全期望公式就是

$$E\xi = \int E(\xi \mid \eta = y) f_\eta(y) \mathrm{d}y.$$

它是全概率公式的推广. 因为对于任意随机事件 A，我们取如下的随机变量

$$\xi = I_A(\omega),$$

即

$$\xi = \begin{cases} 1, & \text{若随机事件 } A \text{ 发生}, \\ 0, & \text{若随机事件 } A \text{ 未发生}. \end{cases}$$

它就成为全概率公式的如下推广形式：

$$P(A)(= EI_A = E(E(I_A \mid \eta))) = \int P(A \mid \eta = y) f_\eta(y) \mathrm{d}y.$$

无论是在理论上，或者是在实用中，这两个公式都是非常有用的.

用概率理论的逼近方法，对于任意非负的或有界的 Borel 函数 $h(x)$ 还可得到下面的结论.

定理 2.2 $E(h(\eta)E(\xi|\eta)) = E(h(\eta)\xi)$.

注意，表达式(E.2)是它的特殊情形，即 $h(y) = I_A(y)$ 的情形.

由于从直观看，$h(\eta)$ 相对于 η 来说，等于是"常数"，所以有下面的结果.

定理 2.3 $E(h(\eta)\xi|\eta) = h(\eta)E(\xi|\eta)$.

如果将视为已知条件的随机变量 η 视为信息源，那么，定理 2.3 说明，相对于信息源为可知的随机变量 $h(\eta)$，在对信息源取条件期望时，可以当作常数处理，这是条件期望的最重要的直观性质，在化简条件期望时极其有用.

例 2.5 设 η 是可能取值为 $y_1, y_2, \cdots, y_j, \cdots$ 的离散随机变量. 对于任意随机变量 ξ，求 $E(\xi|\eta)$.

解 如果一个随机变量 ζ 为 η 可知的，那么它在 $\{\omega: \eta = y_j\}$ $(j = 1, 2, \cdots)$ 上必须取常数值. 所以，随机变量 $E(\xi|\eta)$ 在 $\{\omega: \eta = y_j\}$ 上应取某个常数值 a_j $(j = 1, 2, \cdots)$. 我们通过全期望公式来确定这些常数如下：

$$\begin{aligned} E[\xi \cdot I_{\{y_j\}}(\eta)] &= E[E(\xi \cdot I_{\{y_j\}}(\eta) \mid \eta)] = E[I_{\{y_j\}}(\eta) E(\xi \mid \eta)] \\ &= a_j E[I_{\{y_j\}}(\eta)] = a_j P(\eta = y_j). \end{aligned}$$

故而 $a_j = \dfrac{E[\xi \cdot I_{\{y_j\}}(\eta)]}{P(\eta = y_j)}$. 所以

$$E(\xi \mid \eta) = \sum_j \frac{E[\xi \cdot I_{\{y_j\}}(\eta)]}{P(\eta = y_j)} I_{\{y_j\}}(\eta).$$

特别地，若 ξ 也是离散随机变量，且其可能取值为 $x_1, x_2, \cdots, x_j, \cdots$，则有

$$E[\xi \cdot I_{\{y_j\}}(\eta)] = \sum_i E[\xi \cdot I_{\{x_i\}}(\xi) I_{\{y_j\}}(\eta)]$$

$$= \sum_i x_i P(\xi = x_i, \eta = y_j) \stackrel{\text{def}}{=\!=} \sum_i x_i f(x_i, y_j).$$

因此

$$E(\xi \mid \eta) = \sum_j \frac{\sum_i x_i f(x_i, y_j)}{\sum_i f(x_i, y_j)} I_{\{y_j\}}(\eta).$$

这正是例 2.3 中得到的表达式.

例 2.6 设 A 是随机事件,则其示性函数 $I_A(\omega)$是随机变量. 对于另一个随机变量 ξ 有

$$E(\xi \mid I_A) = \frac{E(\xi \cdot I_A)}{P(A)} I_A(\omega) + \frac{E(\xi \cdot I_{A^c})}{P(A^c)} I_{A^c}(\omega).$$

证明 因为随机变量 $E(\xi|I_A)$为 I_A 可知的,那么它在 A 上应取常数值. 所以 $E(\xi|I_A)$ 在 A 上应取常数值,记此值为 a. 我们通过全期望公式来确定 a 如下:

$$E(\xi \cdot I_A) = E[E(\xi \cdot I_A \mid I_A)] = E[I_A E(\xi \mid I_A)] = aEI_A = aP(A).$$

故而有

$$a = \frac{E(\xi \cdot I_A)}{P(A)}.$$

类似地,$E(\xi|I_A)$在 A^c 上也应取常数,这个值为$\bar{a} = \dfrac{E(\xi \cdot I_{A^c})}{P(A^c)}$.

我们强调,这里的 ξ 可以是离散的,或者是连续型的,还可以是更为一般的随机变量.

我们还容易由定义直接验证下面的结论.

定理 2.4 若随机变量 ξ 与 η 相互独立,则条件期望

$$E(\xi \mid \eta) = E\xi.$$

2.2.3 关于多维随机变量的条件期望

类似地,我们仍然有:随机变量 ξ 关于 d 维随机向量$\boldsymbol{\eta}$ 的条件期望 $E(\xi|\boldsymbol{\eta})$是一个随机变量,它是$\boldsymbol{\eta}$ 的函数(即(E.1)),而且满足表达式(E.2):对于任意 d 维 Borel 集 Λ,恒有

$$E(I_\Lambda(\boldsymbol{\eta}) E(\xi \mid \boldsymbol{\eta})) = E(I_\Lambda(\boldsymbol{\eta})\xi).$$

而类似的结论如下.

定理 2.1′(全期望公式)

$$E[E(\xi \mid \boldsymbol{\eta})] = E\xi.$$

定理 2.2′ $E(h(\boldsymbol{\eta})E(\xi|\boldsymbol{\eta})) = E(h(\boldsymbol{\eta})\xi)$.

定理 2.3′ $E(h(\boldsymbol{\eta})\xi|\boldsymbol{\eta}) = h(\boldsymbol{\eta})E(\xi|\boldsymbol{\eta})$.

定理 2.4′ 若随机向量$\boldsymbol{\xi}$ 与随机向量$\boldsymbol{\eta}$ 相互独立,则

$$E(f(\boldsymbol{\xi}) \mid \boldsymbol{\eta}) = Ef(\boldsymbol{\xi}).$$

下面的关系重要而直观.

定理 2.5

$$E(g(\boldsymbol{\xi},\boldsymbol{\eta})\mid\boldsymbol{\eta}) = [E(g(\boldsymbol{\xi},\boldsymbol{y})\mid\boldsymbol{\eta})]_{\boldsymbol{y}=\boldsymbol{\eta}}.$$

当 $g(x,y)=l(x)h(y)$时,定理显然成立.一般情形的证明,需要用概率论中的典型逼近方法,本书从略.

推论 2.6 当随机向量$\boldsymbol{\xi}$与随机向量$\boldsymbol{\eta}$相互独立时,有

$$E(g(\boldsymbol{\xi},\boldsymbol{\eta})\mid\boldsymbol{\eta}) = [Eg(\boldsymbol{\xi},\boldsymbol{y})]_{\boldsymbol{y}=\boldsymbol{\eta}}.$$

直观地我们还可以接受以下的断言.

定理 2.7(线性性质)

$$E((\alpha\xi+\beta\zeta)\mid\boldsymbol{\eta}) = \alpha E(\xi\mid\boldsymbol{\eta})+\beta E(\zeta\mid\boldsymbol{\eta}).$$

从条件期望的直观含义,我们还有下面的定理.

定理 2.8(平滑性质)

$$E(E(\xi\mid\eta)\mid\eta,\zeta) = E(\xi\mid\eta),$$
$$E(E(\xi\mid\eta,\zeta)\mid\eta) = E(\xi\mid\eta).$$

最后,由于关于$\boldsymbol{\eta}$的条件期望就是在$\boldsymbol{\eta}$确定时的预期,可以证明它有以下的**最佳近似性质**:

定理 2.9(条件期望的二阶矩方法) 若 $Eg(\boldsymbol{\xi})^2<+\infty$,则

$$E[g(\boldsymbol{\xi})-E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})]^2 = \min_{\text{一切Borel函数}h} E[g(\boldsymbol{\xi})-h(\boldsymbol{\eta})]^2.$$

证明 因为对于任意 Borel 函数 h 有

$$\begin{aligned}
&E([g(\boldsymbol{\xi})-E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})][E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta})])\\
=&E[g(\boldsymbol{\xi})(E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta}))]-E[E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})(E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta}))]\\
=&E[g(\boldsymbol{\xi})(E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta}))]-E[E([E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta})]g(\boldsymbol{\xi})\mid\boldsymbol{\eta})]\\
=&E[g(\boldsymbol{\xi})E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})]-E[g(\boldsymbol{\xi})h(\boldsymbol{\eta})]-E([E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})g(\boldsymbol{\xi})]\mid\boldsymbol{\eta})\\
&-E[E([g(\boldsymbol{\xi})h(\boldsymbol{\eta})]\mid\boldsymbol{\eta})]=0,
\end{aligned}$$

所以

$$\begin{aligned}
E[g(\boldsymbol{\xi})-h(\boldsymbol{\eta})]^2 &= E[g(\boldsymbol{\xi})-E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})]^2+E[E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta})]^2\\
&\quad+2E([g(\boldsymbol{\xi})-E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})][E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta})])\\
&= E[g(\boldsymbol{\xi})-E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})]^2+E[E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})-h(\boldsymbol{\eta})]^2\\
&\geqslant E[g(\boldsymbol{\xi})-E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta})]^2.
\end{aligned}$$

这个定理的直观含义是:若

$$z\overset{\text{def}}{=\!=}E(g(\boldsymbol{\xi})\mid\boldsymbol{\eta}),\quad x\overset{\text{def}}{=\!=}g(\boldsymbol{\xi}),$$

$V\overset{\text{def}}{=\!=}\{\boldsymbol{\eta}$的一切 Borel 函数 $h(\boldsymbol{\eta})$生成的闭线性空间(包含线性空间中的序列的均方极限)$\}$,那么,z 是 x 在空间 V 上的投影,记为

$$z=\mathrm{Pro}_V(x).$$

参见图 2.1.

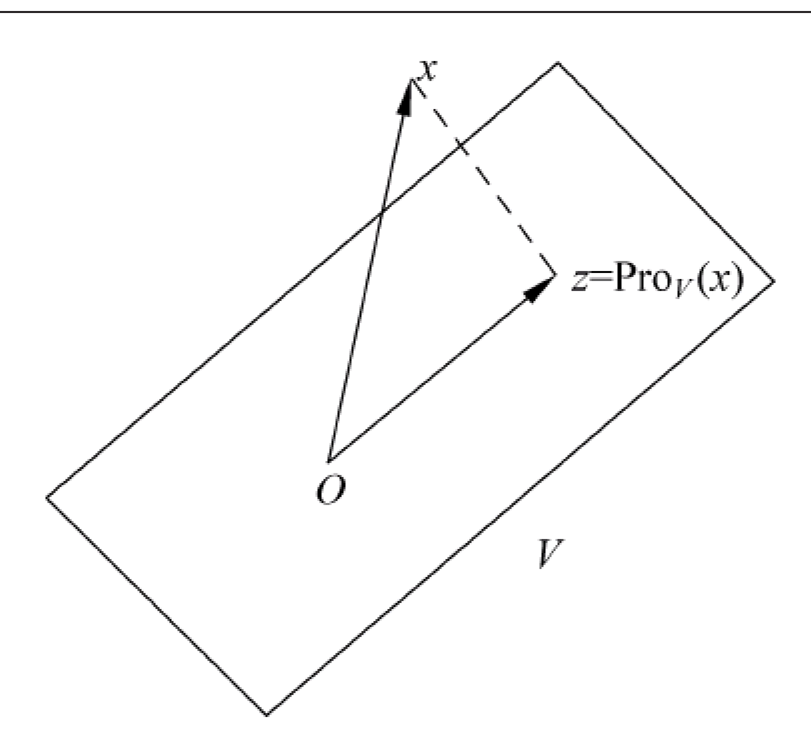


图 2.1　方差有限的随机变量的条件期望理解为投影

2.2.4　方差有限的随机变量对一个随机变量列的条件期望

若 $E\xi^2<+\infty$,则定义

$$E(\xi \mid \eta_k: k=1,2,\cdots) \stackrel{\text{def}}{=\!=} (L^2)\lim_{n\to+\infty} E(\xi \mid \boldsymbol{\eta}^n)\quad \text{(均方收敛的极限)},$$

其中

$$\boldsymbol{\eta}^n \stackrel{\text{def}}{=\!=} (\eta_1,\eta_2,\cdots,\eta_n).$$

2.2.5　期望有限的随机变量对随机变量族的条件期望

将随机变量族 $\eta \stackrel{\text{def}}{=\!=} \{\eta_t: t\geqslant 0\}$ 中任意有限个元素的任意有界连续函数复合的全体所组成的集合记为 $\Phi(\eta)$,包含 $\Phi(\eta)$ 且对收敛性封闭的最小集合记为 $\bar{\Phi}(\eta)$. $\bar{\Phi}(\eta)$ 中的元素是随机变量,称为 **$\boldsymbol{\eta}$-可知的**随机变量.

如果期望有限的(方差可以无限),且某个 η-可知的随机变量 $\hat{\xi}$ 满足:对于任意 η-可知的有界随机变量 ζ,都有

$$E(\xi\zeta)=E(\hat{\xi}\zeta),$$

那么,我们将 $\hat{\xi}$ 定义为 ξ 关于信息 $\{\eta_t, t\geqslant 0\}$ 的条件期望,即

$$E(\xi \mid \eta_t, t\geqslant 0) \stackrel{\text{def}}{=\!=} \hat{\xi}.$$

可以证明,如此定义的条件期望仍满足前面所述的各个相应的定理.

*2.2.6　关于事件体(σ 代数)$\mathscr{F}$的条件期望

一个事件体$\mathscr{F}$表示一个信息源的随机事件的全体.关于此事件体的随机变量,也称为 **$\mathscr{F}$可知的**随机变量.

在不致混淆的情形下,将全体$\mathscr{F}$可知的随机变量也记为$\mathscr{F}$.

如果存在有限期望,而且$\mathscr{F}$可知的随机变量 $\hat{\xi}$,使得对于任意 Φ 中$\mathscr{F}$可知的有界随机

变量 ζ，都满足

$$E(\xi\zeta) = E(\hat{\xi}\zeta),$$

那么，将 $\hat{\xi}$ 称为 ξ **关于事件体 $\mathscr{F}$ 的条件期望**（在不计零概率的差异下，它也是完全确定的），并将它记为 $E(\xi|\mathscr{F})$. 即

$E(\xi|\mathscr{F})$ 是 $\mathscr{F}$ 可知的随机变量，而且对于任意 $\mathscr{F}$ 可知的有界随机变量 ζ 满足：$E(\xi\zeta)=E(E(\xi|\mathscr{F})\zeta)$.

作为特殊情形，当 $\mathscr{F}$ 是 $\{\eta_t: t\geqslant 0\}$ 生成的事件体，意即它是包含所有形如 $\{\omega: \eta_t(\omega)\leqslant x\}$（任意 $t\geqslant 0$，任意实数 x）的事件的最小事件体，则由定义可知

$$E(\xi \mid \mathscr{F}) = E(\xi \mid \eta_t: t \geqslant 0).$$

这与 2.2.5 节中的定义是一致的. 所以此定义是前面所有的定义的推广.

类似地也有

(1) $E(\xi|\mathscr{F})$ 对 ξ 具有线性性质，且对常数 c 有 $E(c|\mathscr{F})=c$.

(2) 若 $\eta\in\mathscr{F}$（即对于任意实数 x，样本点集合 $\{\omega: \eta(\omega)\leqslant x\}\in\mathscr{F}$），则

$$E(\xi\eta \mid \mathscr{F}) = \eta E(\xi \mid \mathscr{F}).$$

即对于信息源为可知的随机变量，在对信息源取条件期望时可以当作常数处理.

(3) 若 ξ 与 σ 代数 $\mathscr{F}$ 相互独立（意即事件体 $\mathscr{F}$ 中的任意事件都与随机变量 ξ 独立），则

$$E(\xi \mid \mathscr{F}) = E\xi.$$

(4)（平滑性质）　若另有一个事件体 $\mathscr{G}$，满足：

$$\mathscr{G}\subset \mathscr{F} \quad （即 \mathscr{G} 比 \mathscr{F} 小）,$$

那么，有

$$E[E(\xi \mid \mathscr{F}) \mid \mathscr{G}) = E(\xi \mid \mathscr{G}),$$
$$E[E(\xi \mid \mathscr{G}) \mid \mathscr{F}) = E(\xi \mid \mathscr{G}).$$

类似地也有最佳近似性质.

例 2.7　若 $\mathscr{F}=\{\varnothing,\Omega\}$. 那么，按可知性的定义，只有常数才是 $\mathscr{F}$ 可知的. 可见 $E(\xi|\mathscr{F})$ 应该是常数. 在等式 $E(\xi\zeta)=E(\hat{\xi}\zeta)$ 中置 $\zeta=1$，便得 $\hat{\xi}=E\xi$，故而

$$E(\xi \mid \mathscr{F}) = E\xi.$$

例 2.8　若 $\mathscr{F}$ 由一个随机事件 A 所生成，我们将它记为

$$\mathscr{F}_A \equiv \{\varnothing, A, A^c \equiv \Omega - A, \Omega\},$$

那么，对其可知的随机变量必须在 A 上取常数，而 $\hat{\xi}$ 在 A 上取的常数值必须为 $\dfrac{E(\hat{\xi}I_A)}{P(A)}$.

同样，$\hat{\xi}$ 在 A^c 上也取常数，且其值必须为 $\dfrac{E(\hat{\xi}I_{A^c})}{P(A^c)}$. 再在等式 $E(\xi\zeta)=E(\hat{\xi}\zeta)$ 中，分别取 $\zeta=I_A(\omega)$ 及 $I_{A^c}(\omega)$ 便得到

$$E(\hat{\xi}I_A) = E(\xi I_A), \quad E(\hat{\xi}I_{A^c}) = E(\xi I_{A^c}).$$

故而

$$
\begin{aligned}
E(\xi \mid \mathscr{F}_A) &= \hat{\xi} = \hat{\xi} I_A(\omega) + \hat{\xi} I_{A^c}(\omega) \\
&= \frac{E(\hat{\xi} I_A)}{P(A)} I_A(\omega) + \frac{E(\hat{\xi} I_{A^c})}{P(A^c)} I_{A^c}(\omega) \\
&= E(\xi \mid I_A) I_A(\omega) + E(\xi \mid I_{A^c}) I_{A^c}(\omega).
\end{aligned}
$$

例 2.9　若$\mathscr{F}$由两个随机事件 A,B 所生成，我们将它记为

$$\mathscr{F}_{A,B} \equiv \{\varnothing, \Omega, A, A^c, B, B^c, AB, AB^c, A^cB, A^cB^c, AB^c \cup A^cB, (AB^c \cup A^cB)^c\}.$$

这时，随机事件 AB,AB^c,A^cB,A^cB^c 构成 Ω 的一个划分（互不相容且其并集为 Ω）. 那么，$\mathscr{F}_{A,B}$可知的随机变量必须在 AB,AB^c,A^cB,A^cB^c 上分别取常数. 从而 $\hat{\xi}$ 在这些随机事件上应分别取常数值 $E(\hat{\xi} I_{AB}),E(\hat{\xi} I_{AB^c}),E(\hat{\xi} I_{A^cB}),E(\hat{\xi} I_{A^cB^c})$. 在等式 $E(\xi\zeta)=E(\hat{\xi}\zeta)$中，分别取 ζ 为 $I_{AB}(\omega)I_{AB^c}(\omega)I_{A^cB}(\omega)I_{A^cB^c}(\omega)$，便得到

$$
\begin{aligned}
E(\xi \mid \mathscr{F}_{A,B}) &= \frac{E(\xi \cdot I_{AB})}{P(AB)} I_{AB}(\omega) + \frac{E(\xi \cdot I_{A^cB})}{P(A^cB)} I_{A^cB}(\omega) \\
&\quad + \frac{E(\xi \cdot I_{AB^c})}{P(AB^c)} I_{AB^c}(\omega) + \frac{E(\xi \cdot I_{A^cB^c})}{P(A^cB^c)} I_{A^cB^c}(\omega) \\
&= E(\xi \mid I_{AB}) I_{AB}(\omega) + E(\xi \mid I_{A^cB}) I_{A^cB}(\omega) \\
&\quad + E(\xi \mid I_{AB^c}) I_{AB^c}(\omega) + E(\xi \mid I_{A^cB^c}) I_{A^cB^c}(\omega).
\end{aligned}
$$

这个例子可以直接推广到由有限个事件生成的事件体$\mathscr{F}$的情形.

注意，在以上的表示式中，ξ 可以是连续型的，可以是更为一般的.

例 2.10　在 η 为离散型随机变量时，可以将 $E(\xi|\eta)$看成 $E(\xi|\mathscr{F})$的特例，具体过程如下：

设随机变量 η 的可能值为 $y_1,y_2,\cdots,y_j,\cdots$，于是事件组

$$A_j \stackrel{\text{def}}{=\!=} \{\omega: \eta(\omega) = y_j\} \quad (j = 1,2,\cdots)$$

构成 Ω 的一个划分. 将它们生成的事件体记为$\mathscr{F}_\eta$，那么，仿照例 2.7 就可以得到

$$
\begin{aligned}
E(\xi \mid \mathscr{F}_\eta) &= \sum_j \frac{E(\xi I_{\{\eta=y_j\}}(\omega))}{P(\eta = y_j)} I_{\{\eta=y_j\}}(\omega) \\
&= \sum_j \frac{E(\xi I_{\{y_j\}}(\eta))}{P(\eta = y_j)} I_{\{y_j\}}(\eta).
\end{aligned}
$$

由例 2.5 可以知道它就是 $E(\xi|\eta)$.

习题 2

1. $X_1,X_2,\cdots,X_n$ 独立同分布，其分布为$\begin{pmatrix} -1 & 1 \\ q & p \end{pmatrix}$. $S_n=X_1+X_2+\cdots+X_n$，$S_0=0$，求 $E(S_3|S_2)$的分布列，并证明 $E(S_{n+1}|S_n)=S_n+(p-q)$.

2. 设 $\xi \sim \mathrm{Exp}_\lambda$，给定 $c>0$，求在 $\xi>c$ 条件下，ξ 的条件分布密度.

3. 若随机变量 X 与 Y 相互独立，且分别服从参数为 λ 与 μ 的指数分布. 定义随机变量

$$Z=\begin{cases}1, & \text{当 } X\leqslant Y,\\ 0, & \text{当 } X>Y.\end{cases}$$

(1) 求条件概率密度 $f_{X|Y}(x|y)$；

(2) 求 Z 的分布列和数学期望.

4. 设 $\eta \sim U[0,1]$，记

$$\eta_n=\sum_{k=0}^{2^n-1}\frac{k}{2^n}I_{\left\{\frac{k}{2^n}\leqslant \eta<\frac{k+1}{2^n}\right\}}.$$

(1) 求 η_n 的分布；

(2) 给定 $\eta_1=1/2$ 下，η_2 的条件分布律.

5. 设随机变量 X 服从[0,1]上的均匀分布，而在随机变量 $X=x$ 条件下，随机变量 Y 的条件分布为 $N(x,x^2)$.

(1) 求 EY，$\mathrm{var}(Y)$，$\mathrm{cov}(X,Y)$.

(2) 证明随机变量 $\dfrac{Y}{X}$ 与 X 独立.

6. 设随机变量 X 服从[0,1]上的均匀分布，而在随机变量 $X=x$ 条件下，随机变量 Y 的条件分布为二项分布 $B(n,x)$.

(1) 求 EY，$\mathrm{var}(Y)$.

(2) 求随机变量 Y 的分布.

7. 设随机变量 X 服从 $\Gamma(\alpha,\lambda)$分布，而在随机变量 $X=x$ 条件下，随机变量 Y 的条件分布为 $\mathrm{Poisson}_x$.

(1) 求 EY，$\mathrm{var}(Y)$.

(2) 求随机变量 Y 的分布.

8. 设随机变量 X 服从 $\Gamma(\alpha,\lambda)$分布，而在随机变量 $X=x$ 条件下，随机变量 N 的条件分布为 $\mathrm{Poisson}_x$，又在随机变量 $N=n$ 的条件下. 随机变量 Y 的条件分布为二项分布 $B(n,p)$.

(1) 证明 $\mathrm{var}(Y)=EX+\mathrm{var}(X)$.

(2) 求随机变量 Y 的分布.

第 3 章　随机徘徊与鞅论浅述

3.1　随机徘徊

3.1.1　简单随机徘徊

定义 3.1　在一个初始随机变量 ξ_0 上，加上一个独立同分布的随机变量列 $\{Y_k, k\geqslant 0\}$ 的部分和，得到的随机变量列 $\{\xi_n\}$，即

$$\xi_n = \xi_0 + \sum_{k=1}^{n} Y_k, \tag{3.1}$$

称为**随机徘徊**，其中 ξ_0 称为**初始位置**.

定义 3.2　如果随机徘徊中的随机变量 Y_k 只取 1 与 -1 两个值，即

$$P(Y_k = 1) = p, \quad P(Y_k = -1) = q \overset{\text{def}}{=\!=} 1 - p,$$

而且 ξ_0 只取整数，那么，$\{\xi_n\}$ 称为**简单随机徘徊**. 这里 ξ_n 表示一个粒子从随机位置 ξ_0 出发，每次分别以概率 p 或 q，向右或向左走一格. $p=\dfrac{1}{2}$ 的情况称为**对称简单随机徘徊**. 易见，对于任意整数 i, j 及 $\xi_0=0$ 有

$$p_{ij} \overset{\text{def}}{=\!=} P(\xi_{n+1} = j \mid \xi_n = i) = \begin{cases} p, & j = i+1, \\ q, & j = i-1, \\ 0, & \text{其他}, \end{cases}$$

$$P(\xi_n = j) = \begin{cases} C_n^{\frac{n+j}{2}} p^{\frac{n+j}{2}} (1-p)^{\frac{n-j}{2}}, & \text{若 } n \geqslant |j|, \text{且 } n \text{ 与 } j \text{ 奇偶同}, \\ 0. & \text{其他情形}, \end{cases}$$

此外对 $\xi_0 = i$ 还有

$$\begin{aligned} p_{ij}^{(n)} &\overset{\text{def}}{=\!=} P(\xi_n = j \mid \xi_0 = i) \\ &= \begin{cases} C_n^{\frac{n+j-i}{2}} p^{\frac{n+j-i}{2}} (1-p)^{\frac{n-j+i}{2}}, & \text{若 } n \geqslant |j-i|, \text{且 } n \text{ 与 } j-i \text{ 奇偶同}, \\ 0, & \text{其他情形}. \end{cases} \end{aligned}$$

对称简单随机徘徊是**常返的**，即对于任意整数 i 有

$$P(\text{存在 } n > 1 \text{ 使 } \xi_n = i \mid \xi_0 = i) = 1.$$

(1) 简单随机徘徊的联合分布

对于 $m>n$ 有

$$P(\xi_n = r, \xi_m = s) = P(\xi_n = r, \xi_m - \xi_n = s - r)$$
$$= \mathrm{C}_n^{\frac{n+r}{2}} p^{\frac{n+r}{2}} (1-p)^{\frac{n-r}{2}} \cdot \mathrm{C}_{m-n}^{\frac{m-n+s-r}{2}} p^{\frac{m-n+s-r}{2}} (1-p)^{\frac{m-n-s+r}{2}}.$$

对于 $n_1 < n_2 < \cdots < n_k$，类似地有

$$P(\xi_{n_1} = s_1, \xi_{n_2} = s_2, \cdots, \xi_{n_k} = s_k)$$
$$= \prod_{l=1}^{k} \mathrm{C}_{m_l}^{r_l} p^{r_l} (1-p)^{m_l - r_l} \quad \left(m_l = n_l - n_{l-1}, r_l = \frac{1}{2}(n_l - n_{l-1} + s_l - s_{l-1}) \right).$$

（2）简单随机徘徊的均值、方差以及协方差.

初始位置为固定值 x 的简单随机徘徊记为

$$\xi_n^x \stackrel{\text{def}}{=\!=} x + Y_1 + Y_2 + \cdots + Y_n. \tag{3.2}$$

我们有

$$E\xi_n^x = x + nEY_i = x + n(p-q),$$
$$\mathrm{var}(\xi_n^x) = n\mathrm{var}(Y_i) = 4npq.$$

特别地，对于从原点出发的对称简单随机徘徊有

$$E\xi_n^0 = 0, \quad \mathrm{var}(\xi_n^0) = n.$$

当 $n < m$ 时，由于

$$E(\xi_n^x \xi_m^x) = E[(x + Y_1 + Y_2 + \cdots + Y_n)(x + Y_1 + Y_2 + \cdots + Y_m)]$$
$$= x^2 + x(m+n)(p-q) + \sum_{i=1}^{n} \sum_{j=1}^{m} E(Y_i Y_j)$$
$$= x^2 + x(m+n)(p-q) + \sum_{i=1}^{n} E(Y_i^2) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{m} E(Y_i)E(Y_j)$$
$$= x^2 + x(m+n)(p-q) + n + [nm - n](p-q)^2.$$

故协方差

$$\mathrm{cov}(\xi_n^x, \xi_m^x) = E(\xi_n^x \xi_m^x) - E\xi_n^x E\xi_m^x$$
$$= x^2 + x(m+n)(p-q) + n + n(m-1)(p-q)^2$$
$$- [x + n(p-q)][x + m(p-q)]$$
$$= n - n(p-q)^2 = 4npq.$$

所以，对于任意的 n, m，就有

$$\mathrm{cov}(\xi_n^x, \xi_m^x) = 4(n \wedge m)pq,$$

其中

$$n \wedge m \stackrel{\text{def}}{=\!=} \min\{n, m\}. \tag{3.3}$$

3.1.2 博弈输光模型

设甲乙两人博采，每次 1 元，直至其中一人输光为止. 设在时刻 0 甲有初始资金 a 元，

乙有初始资金 b 元. 假定每次博采时,甲赢的概率是 p,而乙赢的概率是 $q=1-p$,而且各局的进行是彼此独立的,规定不欠不借,博采一直到甲乙中有一人输光才结束. 求甲输光的概率 p_a 和乙输光的概率 q_b.

由全概率公式,甲输光的概率 p_a 应满足如下的递推关系

$$p_a = pp_{a+1} + qp_{a-1}.$$

这是一个常系数的二阶差分方程,并且满足两个边界条件:

$$p_0 = 1,\quad p_{a+b} = 0.$$

对 a 归纳地可以求得这个差分方程的解为

$$p_a = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}} \quad \left(p \neq \frac{1}{2}\right),$$

$$p_a = \frac{b}{a+b} \quad \left(p = \frac{1}{2}\right).$$

从对称性同样可得到,乙输光的概率为

$$q_b = \frac{\left(\frac{p}{q}\right)^b - \left(\frac{p}{q}\right)^{a+b}}{1-\left(\frac{p}{q}\right)^{a+b}} \quad \left(p \neq \frac{1}{2}\right),$$

$$q_b = \frac{a}{a+b} \quad \left(p = \frac{1}{2}\right).$$

由此可见有

$$p_a + q_b = 1.$$

3.2　鞅列浅述

3.2.1　鞅列与常见的例子

定义 3.3　一个随机序列 $\{\eta_n: n\geqslant 0\}$ 可以作为一系列"历史事件"的参照信息. 我们将它简单地记为

$$\boldsymbol{\eta}_n \xlongequal{\text{def}} \{\eta_0, \eta_1, \cdots, \eta_n\}.$$

于是,随机变量集合 $\bar{\Phi}(\boldsymbol{\eta}_n)$(记号参见 2.2.5 节)代表 n-前有用信息,称为 ***n* 前的历史**. 随机变量 ξ 称为 $\boldsymbol{\eta}_n$ **可知的**,如果 $\xi\in\bar{\Phi}(\boldsymbol{\eta}_n)$.

定义 3.4　随机序列 $\{\xi_n: n\geqslant 0\}$ 称为 (η_n) **可知的**,或者 (η_n) **适应的**,如果对于任意 m, ξ_m 都是 $\boldsymbol{\eta}_m$ 可知的.

注　请务必分清本书中所使用的记号的差别, η_n 是指一个随机变量,代表"现在的"

信息.而黑体的$\boldsymbol{\eta}_m$则表示一段随机变量列$\{\eta_m: m\leqslant n\}$,代表"现在及以前的历史"信息.

在讨论鞅列,或以后的鞅的时候,凡是论及两个随机变量的相等,一律理解为概率为1地相等.也就是说,我们从不区分两个概率为1地相等的随机变量.

定义 3.5　期望值有限的随机序列$\{\eta_n: n\geqslant 0\}$,如果满足:对于任意m都有

$$E(\eta_{n+m} \mid \boldsymbol{\eta}_n) = \eta_n,$$

则称为**鞅列**.另一个期望值有限的随机序列$\{\xi_n: n\geqslant 0\}$,如果满足$\{\xi_n\}$是$(\boldsymbol{\eta}_n)$可知的,且对于任意m都有

$$E(\xi_{n+m} \mid \boldsymbol{\eta}_n) = \xi_n,$$

则称为(η_n)鞅列.由条件期望的性质容易看出,若ξ_n是(η_n)鞅列,则它也是鞅列.

如果上面式子中的"="分别改为"≥"或"≤",则对应的随机序列分别称为**下鞅列**,或**上鞅列**.

命题 3.1　下面两个叙述彼此等价:

(1) 对于任意$n,m,E(\xi_{n+m}|\boldsymbol{\eta}_n)=\xi_n$.

(2) 对于任意$n,E(\xi_{n+1}|\boldsymbol{\eta}_n)=\xi_n$,即$\xi_n$为$(\eta_n)$鞅列.

证明　对m作归纳法,由条件期望的平滑性质得到

$$E(\xi_{n+m+1} \mid \boldsymbol{\eta}_n) = E(E(\xi_{n+m+1} \mid \boldsymbol{\eta}_n) \mid \boldsymbol{\eta}_n) = E(\xi_{n+1} \mid \boldsymbol{\eta}_n) = \xi_n.$$

我们来分析鞅列这个模型的直观含义.设η_n是鞅列.假定某人在时刻n持有本金η_n,并参加对弈.那么,在下一时刻$n+1$,他的资金η_{n+1}是随机的.但是,应用条件期望的平滑性质,在"最近"的历史资金η_n已知条件下,它在下一时刻$n+1$的条件期望$E(\eta_{n+1}|\boldsymbol{\eta}_n)$是

$$E(\eta_{n+1} \mid \boldsymbol{\eta}_n) = E(E(\eta_{n+1} \mid \boldsymbol{\eta}_n) \mid \boldsymbol{\eta}_n) = E(\eta_{n+1} \mid \boldsymbol{\eta}_n) = \eta_n.$$

这说明在η_n已知的条件下,下一时刻$n+1$的收益的条件平均与现有的本金η_n恰好相等.故而这样的博弈(平均地)是公平的.因此,鞅列表示一个平均趋势公平发展的博弈的本金变化的随机序列.而下鞅列则表示按平均趋势必赢的博弈.上鞅列则表示按平均趋势必输的博弈.

例 3.1(随机徘徊)　$\eta_n=a+\xi_1+\xi_2+\cdots+\xi_n$,$\{\xi_i\}$是独立同分布随机列,且$E\xi_i=0$.于是$\eta_n$是随机徘徊.按定义容易验证它是鞅列.

证明

$$E(\eta_{n+1} \mid \boldsymbol{\eta}_n) = \eta_n + E(\xi_{n+1} \mid \boldsymbol{\eta}_n) = \eta_n.$$

例 3.2　$\xi_n \overset{\text{def}}{=\!=} E(\xi|\boldsymbol{\eta}_n)$是$(\eta_n)$鞅列.

例 3.3　如果随机列$\{\xi_n\}$为(η_n)可知的,定义

$$\zeta_n = \xi_n - \sum_{k=0}^{n-1} [E(\xi_{k+1} \mid \boldsymbol{\eta}_k) - \xi_k],$$

则$\{\zeta_n\}$为(η_n)鞅列.ζ_n是由ξ_n相继地减去它与鞅列的差的波动而得到的,直观地看,它应该就是鞅列.事实上

$$E(\zeta_{n+1} \mid \boldsymbol{\eta}_n) = E(\xi_{n+1} \mid \boldsymbol{\eta}_n) - \sum_{k=0}^{n} (E[E(\xi_{k+1} \mid \boldsymbol{\eta}_k) \mid \boldsymbol{\eta}_n] - E(\xi_k \mid \boldsymbol{\eta}_n))$$
$$= E(\xi_{n+1} \mid \boldsymbol{\eta}_n) - \sum_{k=0}^{n} (E(\xi_{k+1} \mid \boldsymbol{\eta}_k) - \xi_k)$$
$$= \xi_n - \sum_{k=0}^{n-1} (E(\xi_{k+1} \mid \boldsymbol{\eta}_k) - \xi_k) = \zeta_n.$$

例 3.4(随机徘徊的平方可积鞅列)　在例 3.1 中,如果还有 $E\xi_i^2 = \sigma^2 < +\infty$,我们定义

$$\zeta_n = \eta_n^2 - \mathrm{var}(\eta_n) = \eta_n^2 - n\sigma^2,$$

那么,$\{\zeta_n : n \geqslant 1\}$为$(\xi_n)$鞅列.

证明　记

$$\boldsymbol{\xi}_n = \{\xi_0, \xi_1, \cdots, \xi_n\}.$$

那么,η_n^2 为(ξ_n)可知的. 因为 ξ_{k+1} 与 ξ_k 独立,所以

$$E\xi_n = 0, \quad \mathrm{var}(\eta_n) = n\sigma^2,$$
$$E(\xi_{k+1} \mid \boldsymbol{\xi}_k) = E\xi_{k+1} = 0, \quad E(\xi_{k+1}^2 \mid \boldsymbol{\xi}_k) = E\xi_{k+1}^2 = \sigma^2.$$

于是

$$E(\eta_{k+1}^2 \mid \boldsymbol{\xi}_k) - \eta_k^2 = E((\eta_k + \xi_{k+1})^2 \mid \boldsymbol{\xi}_k) - \eta_k^2$$
$$= [\eta_k^2 + 2E((\eta_k \xi_{k+1}) \mid \boldsymbol{\xi}_k) + \sigma^2] - \eta_k^2 = 2\eta_k E(\xi_{k+1} \mid \boldsymbol{\xi}_k) + \sigma^2 = \sigma^2.$$

由此得到

$$E(\zeta_{k+1} \mid \boldsymbol{\xi}_k) = E(\eta_{k+1}^2 \mid \boldsymbol{\xi}_k) - (k+1)\sigma^2 = \eta_k^2 - k\sigma^2 = \zeta_k.$$

故而$\{\zeta_n\}$是(ξ_n)鞅列.

例 3.5(随机利率)　设$\{\xi_n\}$是(η_n)可知的非负随机序列. 对$\boldsymbol{\eta}_n = \{\eta_0, \eta_1, \cdots, \eta_n\}$,若随机利率序列$\{\delta_n\}$也是$(\eta_n)$可知的,且满足条件连续利率增长关系

$$E(\xi_{n+1} \mid \boldsymbol{\eta}_n) = \mathrm{e}^{\delta_n} \xi_n.$$

将时刻 n 的资金 ξ_n 在开始时刻的折现的价值记为 ζ_n,即

$$\zeta_n = \mathrm{e}^{-(\delta_0 + \delta_1 + \cdots + \delta_{n-1})} \xi_n.$$

那么$\{\zeta_n\}$为(η_n)鞅列.

证明

$$E(\zeta_{n+1} \mid \boldsymbol{\eta}_n) = \mathrm{e}^{-(\delta_0 + \delta_1 + \cdots + \delta_n)} E(\xi_{n+1} \mid \boldsymbol{\eta}_n) = \zeta_n.$$

例 3.6(指数鞅列)　在例 3.1 中令

$$\zeta_n = \prod_{k=1}^{n} \frac{\mathrm{e}^{s\xi_k}}{E\,\mathrm{e}^{s\xi_k}},$$

则$\{\zeta_n\}$为(ξ_n)鞅列.

证明　记$\boldsymbol{\xi}_n = \{\xi_0, \xi_1, \cdots, \xi_n\}$. 由 ξ_{n+1} 与 ξ_n 的独立性得到

$$E(\zeta_{n+1} \mid \boldsymbol{\xi}_n) = \zeta_n \frac{1}{Ee^{s\xi_{n+1}}} E(e^{s\xi_{n+1}} \mid \boldsymbol{\xi}_n) = \zeta_n.$$

例 3.7(分支鞅列) 设开始时刻有一个细胞(称为第 0 代),它在下一个整数时刻分裂为第一代的细胞后随即死亡,各个新细胞又按相同的规律独立地在下一个整数时刻再进行同样方式的分裂.将第 n 代的第 k 个细胞分裂的个数记为 $\eta_{n,k}$,它是一个随机变量,对于不全相同的 k,n 而言,$\eta_{n,k}$ 们是独立同分布的,假定它们有有限的期望 μ.将经过第 n 代分裂后得到的细胞总数记为 ξ_n.于是

$$\begin{aligned} &\xi_0 = 1, \\ &\vdots \\ &\xi_0 = 1, \quad \xi_n = \eta_{n,1} + \eta_{n,2} + \cdots + \eta_{n,\xi_{n-1}}. \end{aligned}$$

那么 $\left\{\frac{\xi_n}{\mu^n}\right\}$ 是 $\{\xi_n\}$ 鞅列.

证明 由各代之间以及同一代的各个个体间的独立性,有

$$\begin{aligned} E\left(\frac{\xi_n}{\mu^n} \mid \xi_0, \xi_1, \cdots, \xi_{n-1}\right) &= E\left(\frac{\eta_{n,1} + \eta_{n,2} + \cdots + \eta_{n,\xi_{n-1}}}{\mu^n} \mid \xi_0, \xi_1, \cdots, \xi_{n-1}\right) \\ &= E\left(\frac{\eta_{n,1} + \eta_{n,2} + \cdots + \eta_{n,\xi_{n-1}}}{\mu^n} \mid \xi_{n-1}\right) \\ &= \frac{1}{\mu^n} \sum_{k=1}^{\xi_{n-1}} E\eta_{n,k} = \frac{\xi_{n-1}\mu}{\mu^n} = \frac{\xi_{n-1}}{\mu^{n-1}}. \end{aligned}$$

可见结论成立.

例 3.8(似然比序列) 设 $\{\xi_n\}$ 为独立同分布列,具有分布密度 f.如果 g 为另一个随机变量的分布密度,那么样本序列关于此两个分布密度的似然比

$$\eta_n \overset{\text{def}}{=\!=} \prod_{k=1}^{n} \frac{g(\xi_k)}{f(\xi_k)}$$

是 (ξ_n) 鞅列.

证明 因为

$$E\left(\frac{g(\xi_n)}{f(\xi_n)}\right) = \int \frac{g(x)}{f(x)} f(x)\,dx = 1,$$

所以

$$E(\eta_n \mid \xi_1, \xi_2, \cdots, \xi_{n-1}) = \eta_{n-1} E\left(\frac{g(\xi_n)}{f(\xi_n)} \mid \xi_1, \xi_2, \cdots, \xi_{n-1}\right) = \eta_{n-1} E\left(\frac{f(\xi_n)}{g(\xi_n)}\right) = \eta_{n-1}.$$

3.2.2 鞅列的停时与选样定理

(1) 停时

命题 3.2 设 $\{\xi_n\}$ 为 (η_n) 鞅列,则 $E\xi_n = E\xi_0$.

证明 用全期望公式便得到

$$E\xi_{n+1} = E[E(\xi_{n+1} \mid \boldsymbol{\eta}_n)] = E\xi_n.$$

定义 3.6 一个可以取值$+\infty$的非负整值随机时刻 τ，称为**关于$\{\eta_n\}$的停时**(简称**$\{\eta_n\}$停时**)，如果 τ 在 n 时发生与否，只依赖于$\{\eta_n\}$的历史记录$\boldsymbol{\eta}_n=\{\eta_0,\eta_1,\cdots,\eta_n\}$，即复合随机变量 $I_{\{n\}}(\tau)$是$\boldsymbol{\eta}_n$ 可知的，其中

$$I_{\{n\}}(t) = \begin{cases} 1, & t = n, \\ 0, & t \neq n. \end{cases}$$

停时的直观含义 博弈的参与者，常常根据他事先设计好的规则，在每一个时刻决定继续还是退出(典型的规则有，输完本钱时退出，输赢超过本钱的 80%时退出等). 这些事先设计好的时刻，实际上都是随机的. 所以，根据这些事先设计好的随机时刻退出时，基本事件的如下集合：{在时刻 n 退出}就是一个随机事件，由博弈者在时刻 n 与时刻 n 以前的累计得失的历史记录可以完全确定. 因此，这些设计好的随机时刻都是停时.

从停时的直观含义，不难建立以下的命题.

命题 3.3 常数 n 是$\{\eta_n\}$停时；又如果 τ,σ 都是$\{\eta_n\}$停时，$n\geqslant 0$，则

$$\tau \wedge \sigma = \min\{\tau,\sigma\}$$

是$\{\eta_n\}$停时，因此 $\tau\wedge n$ 也是停时.

(2) 随机序列$\{\xi_n\}$在停时 τ 上的取值

在作一次试验得到一个基本事件 ω 后，随机序列$\{\xi_n\}$就相应地得到了其采样(数值)$\{\xi_n(\omega)\}$. 假定 τ 是一个只取有限值的停时，那么 $\tau=\tau(\omega)$就是该基本事件 ω 确定后的一个确切时刻. 我们可以自然地将**随机序列$\{\boldsymbol{\xi}_n\}$在停时$\boldsymbol{\tau}$ 上的取值**理解为

$$\xi_\tau \overset{\text{def}}{=\!=} \xi_{\tau(\omega)}(\omega).$$

(3) 选样定理(option theorem)

本书中将介绍的选样定理的形式，是鞅论中著名的 Doob 停止定理的较为简单的部分. 它说明对于一个公平的博弈，即对于一个公平博弈，在某一类较为规范的停时上的受益，仍然体现出公平性，即如果博弈者在这类停时上退出，其所得关于不同的 ω 取的平均，与博弈开始时的平均资金仍然是一样的.

如上所述，典型的停时是参加博弈的人给自己预先制定的退出博弈的时间 τ. 例如，他的资金达到 1000 元的时刻(若在时刻 n，他有 ξ_n 元，那么，τ 就是 ξ_n 首达 1000 的时刻. 如果对于某种情况(即某个基本事件 ω)，他的资金永远达不到 1000 元，那么，此时的 $\tau(\omega)$就是$+\infty$. 这就同时说明了对于一般的停时，取值$+\infty$是完全可能的).

用一个鞅列所描述的博弈是公平的，即对于任意 n 均有 $E\xi_n=E\xi_0$. 那么在任意停时 τ 上是否也有 $E\xi_\tau=E\xi_0$ 呢？

这个结论一般并不一定成立. 例如，如果$\{\xi_n\}$是满足 $\xi_0=1$ 的对称简单随机徘徊，若

停时 τ 为首达 2 的时刻. 由于对称简单随机徘徊$\{\xi_n\}$可以取负值,这就相当于容许参加博弈的人欠钱. 然而,由于每次输赢的概率都是$\frac{1}{2}$,就可以通过直接计算验证停时 τ 是有限的,即 $P(\tau<+\infty)=1$. 注意,按定义 $\xi_\tau\equiv 2$,我们有 $E\xi_0=1, E\xi_\tau=2$. 即博弈人在时刻 τ 停止的策略是稳赢 1 元的策略,因此它显然不是公平策略,这里产生不公平性的原因在于:让参加博弈的人不加限制地欠钱是一个对于他过分倾斜的有利的条件,这对于他的博弈对手是很不公平的. 由此看来,要满足 $E\xi_\tau=E\xi_0$,就必须对停时 τ 有所限制. 下面的定理给出了常见的充分条件,直观地就是在停时上平均出现的输赢是公平的.

定理 3.4(选样定理) 设 $\xi_n(n\geqslant 0)$是鞅列. 而 τ 是一个取有限值的$\{\xi_n\}$停时,而且满足

(1) τ 是有界停时.

或更一般地

(2) ξ_τ 的期望有限,且 $\lim\limits_{n\to+\infty} E(|\xi_n|\cdot I_{\{\tau>n\}})=0$.

那么,有

$$E\xi_\tau = E\xi_0.$$

这里的假定 $\lim\limits_{n\to+\infty} E(|\xi_n|\cdot I_{\{\tau>n\}})=0$,完全是一个纯粹技术性的数学条件.

证明 先假定 τ 有界,即假定 $\tau\leqslant m$. 因为 τ 是$\{\xi_n\}$停时,所以随机事件$\{\tau\geqslant k\}=\{\tau\leqslant k-1\}^c$ 是由 $\xi_{k-1},\cdots,\xi_0$ 的信息确定的,故 $I_{\tau\geqslant k}$是$(\xi_{k-1},\cdots,\xi_0)$可知的. 于是

$$E[(\xi_k-\xi_{k-1})I_{\tau\geqslant k}\mid \xi_{k-1},\cdots,\xi_0] = I_{\tau\geqslant k}E[(\xi_k-\xi_{k-1})\mid \xi_{k-1},\cdots,\xi_0] = 0.$$

由此得到

$$\begin{aligned} E(\xi_\tau-\xi_0) &= \sum_{j=1}^m E[(\xi_j-\xi_0)I_{\tau=j}] = \sum_{j=1}^m E\Big[\sum_{k=1}^j (\xi_k-\xi_{k-1})I_{\tau=j}\Big] \\ &= \sum_{k=1}^m E\Big[(\xi_k-\xi_{k-1})\sum_{j=k}^m I_{\tau=j}\Big] = \sum_{k=1}^m E[(\xi_k-\xi_{k-1})I_{\tau\geqslant k}] = 0. \end{aligned}$$

对于一般的 τ. 由于 $\tau_m=\tau\wedge m$ 也是$\{\xi_n\}$停时,而且是有界的,故而有 $E\xi_{\tau_m}=E\xi_0$. 从而可以利用假定推出

$$\begin{aligned} |E\xi_\tau-E\xi_{\tau_m}| &= |E[(\xi_\tau-\xi_m)I_{\tau\geqslant m}]\leqslant |E(\xi_\tau I_{\tau\geqslant m})|+|E(\xi_m I_{\tau\geqslant m})| \\ &\leqslant E(|\xi_\tau|I_{\tau\geqslant m})+E(|\xi_m|I_{\tau\geqslant m})\xrightarrow{m\to+\infty} 0, \end{aligned}$$

其中前一项趋于 0 是由于假定 ξ_τ 的期望有限,而后一项趋于 0 是因为假定了 $\lim\limits_{n\to+\infty} E(|\xi_n|\cdot I_{\{\tau>n\}})=0$. 最后,就得到 $E\xi_\tau=E\xi_0$.

选样定理告诉我们,对于一个按趋势平均公平的博弈,它在"公平的停时"上也反映了按趋势平均公平的性质.

3.2.3 鞅列的选样定理的应用

鞅列方法与选样定理在应用中的典型例子是博弈输光问题.

例 3.9(输光问题再访) 我们用鞅列的选样定理,重新求甲输光的概率 p_a 与乙输光的概率 q_b,并求此种博弈的平均持续时间.

将在时刻 n 时甲拥有的资金记为 ξ_n. 输光问题实际上是一个简单随机徘徊,有两个吸收壁(含义为:随机徘徊一旦到了吸收壁,以后就停止在该处):0 与 $a+b$(从甲的立场看,甲输光对应于资金停止在 0 上,即被 0 吸收,而乙输光对应于被 $a+b$ 吸收). 需要计算的是被其中的一个吸收的概率. 令$\{\xi_n\}$为简单随机徘徊:

$$\xi_n = a + Y_1 + Y_2 + \cdots + Y_n, \{Y_n: n \geqslant 0\} \text{ 独立同分布,且 } Y_n \sim \begin{pmatrix} -1 & 1 \\ q & p \end{pmatrix}.$$

令

$$\tau = \inf\{n: \xi_n = 0 \text{ 或 } \xi_n = a + b\} \quad (\inf\varnothing(\text{空集}) \overset{\text{def}}{=\!=} +\infty).$$

于是

$$p_a = P(\xi_\tau = 0), \quad q_b = P(\xi_\tau = a + b), \quad p_a + q_b = 1.$$

为了利用鞅列方法计算 p_a,我们区分如下两种情形:

(1) 对称情形:$p=\dfrac{1}{2}$,$\{\xi_n\}$是鞅列.

事实上

$$E(\xi_{n+1} \mid \xi_n, \cdots, \xi_0) = p(\xi_n + 1) + q(\xi_n - 1) = \xi_n,$$

所以$\{\xi_n\}$是鞅列. 这时从直观可以验证 $\tau<+\infty$. 所以

$$E(|\xi_n| I_{\{\tau>n\}}) \leqslant (a+b)EI_{\{\tau>n\}} = (a+b)P(\tau > n) \to 0.$$

从而由定理 3.4 推出

$$E\xi_\tau = E\xi_0 = a.$$

另一方面

$$E\xi_\tau = 0 \cdot P(\xi_\tau = 0) + (a+b)P(\xi_\tau = a+b) = (a+b)q_b.$$

这说明

$$q_b = \frac{a}{a+b},$$

这是乙输光的概率. 进而

$$p_a = 1 - q_b = \frac{b}{a+b}.$$

我们继续求此时博弈的平均持续时间 $E(\tau|\xi_0=a)$. 对此将利用其平方可积鞅列:

$$\eta_n \equiv \xi_n^2 - n\mathrm{var}(Y_1) = \xi_n^2 - n.$$

这是例 3.3 的特殊情形. 但是,因为这种情形非常简单,我们不妨直接验证如下:

$$E(\eta_{n+1} \mid \xi_n,\cdots,\xi_0) = E(\xi_{n+1}^2 \mid \xi_n,\cdots,\xi_0) - (n+1)$$
$$= p(\xi_n+1)^2 + q(\xi_n-1)^2 - (n+1) = \xi_n^2 - n = \eta_n,$$

这就验证了$\{\eta_n\}$是(ξ_n)鞅列. 对于$\tau=\inf\{n: \xi_n=0$ 或 $\xi_n=a+b\}$,直观地,有$E\tau<+\infty$(事实上,假定ζ_n是以 0 与 $a+b$ 为反射壁的反射随机徘徊,那么ζ_n是正常返的. 记$T=\inf\{n: \zeta_n=0$ 或 $\zeta_n=a+b\}$,那么,$ET<+\infty$. 故而$E\tau<ET<+\infty$). 由

$$|\eta_n| \leqslant (a+b)^2 + n \quad (\tau<+\infty)$$

及

$$E(nI_{\tau>n}) < E(\tau I_{\tau>n}) \to 0 \quad (n\to+\infty)$$

推出

$$\lim_{n\to+\infty} E(|\eta_n| I_{\{\tau>n\}}) = 0.$$

用定理 3.4 就得到

$$E\eta_\tau = E\eta_0 = a^2.$$

即

$$a^2 = E\xi_\tau^2 - E\tau = 0 \cdot p_a + (a+b)^2 q_b - E\tau = (a+b)a - E\tau.$$

由此得到博弈的平均持续时间

$$E\tau = ab.$$

(2) $p>q$ 情形. 记 $\mu=p-q>0$.

此时 $0<\dfrac{q}{p}<1$. 这时是不公平的博弈,甲方赢的概率大. 为了计算 p_a, q_b,定义

$$\zeta_n = \left(\frac{q}{p}\right)^{\xi_n}.$$

我们验证$\{\zeta_n\}$是(ξ_n)鞅. 事实上

$$E(\zeta_{n+1} \mid \xi_n,\cdots,\xi_0) = E\left[\left(\frac{q}{p}\right)^{\xi_{n+1}} \mid \xi_n,\cdots,\xi_0\right] = p\left(\frac{q}{p}\right)^{(\xi_n+1)} + q\left(\frac{q}{p}\right)^{(\xi_n-1)}$$
$$= \left(\frac{q}{p}\right)^{\xi_n} = \zeta_n.$$

类似地,对于$\tau=\inf\{n: \xi_n=0$ 或 $\xi_n=a+b\}$,也有 $\lim\limits_{n\to+\infty} E(|\zeta_n| I_{\{\tau>n\}})=0$. 再用定理 3.4,就得到

$$E\zeta_\tau = E\zeta_0 = \left(\frac{q}{p}\right)^a.$$

另一方面,按定义我们还有

$$E\zeta_\tau = \left(\frac{q}{p}\right)^0 p_a + \left(\frac{q}{p}\right)^{a+b}(1-p_a).$$

把它代入前一式,就可以解出

$$p_a = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}, \quad q_b = \frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^{a+b}}. \tag{*}$$

注 类似地，如果将简单随机徘徊的初始出发点改为 $\xi_0=0$，并定义

$$\tau[-a,b] = \inf\{\xi_n = -a \text{ 或 } \xi_n = b\},$$

p_a, q_b 如(*)式所示，那么仍有

$$P(\xi_{\tau[-a,b]} = -a) = p_a, \quad P(\xi_{\tau[-a,b]} = b) = q_b.$$

为了求此不公平博弈的平均持续时间，我们定义

$$X_n = \xi_n - n\mu.$$

它满足

$$\begin{aligned} E(X_{n+1} \mid \xi_n, \cdots, \xi_0) &= E(\xi_{n+1} \mid \xi_n, \cdots, \xi_0) - (n+1)\mu \\ &= p(\xi_n+1) + q(\xi_n-1) - (n+1)\mu = X_n, \end{aligned}$$

即 $\{X_n\}$ 是 (ξ_n) 鞅列. 类似地，可以验证 $\lim\limits_{n\to+\infty} E(|X_n| I_{\{\tau>n\}})=0$. 同样，由定理 3.4 推出 $EX_\tau = EX_0$，即

$$-ap_a + bq_b - \mu E\tau = 0.$$

故而有

$$E\tau = \frac{1}{\mu}(bq_b - ap_a).$$

下面我们假定 $\xi_0=0$. 记 ξ_n 首达 1 的时刻为 τ_1. 我们来计算 $E\tau_1$.

$\{\xi_n\}$停时 τ_1 并不"规范"，即它并不满足定理 3.4 的假设条件，因此，$E\tau_1$ 不能如上面那样直接求，而需要通过极限过渡. 为此令

$$\tau_{[-N,1]} = \inf\{\xi_n = -N \text{ 或 } \xi_n = 1\},$$

它也是$\{\xi_n\}$停时. 可以验证 $\lim\limits_{n\to+\infty} E(|X_n| I_{\{\tau_{[-N,1]}>n\}})=0$. 于是由定理 3.4 得到

$$EX_{\tau_{[-N,1]}} = EX_0,$$

即

$$-Np_N + 1\cdot(1-p_N) - \mu E\tau_{[-N,1]} = 0,$$

其中

$$p_N \equiv P(X_{\tau_{[-N,1]}} = -N).$$

可见

$$E\tau_{[-N,1]} = \frac{1}{\mu}[1-(N+1)p_N] = \frac{1}{\mu}\left[1-(N+1)\frac{\left(\frac{q}{p}\right)^N - \left(\frac{q}{p}\right)^{N+1}}{1-\left(\frac{q}{p}\right)^{N+1}}\right].$$

所以，有

$$E\tau_1 = \lim_{N\to+\infty} E\tau_{[-N,1]} = \frac{1}{\mu} = \frac{1}{p-q}.$$

(3) $p<q$ 情形，与(2)完全类似.

可见，选样定理对于研究对称简单随机徘徊的首达时，是一个重要的数学工具.

* 3.2.4 一般 Doob 停止定理的叙述

一般的 Doob 停止定理，虽然其直观含义仍然是很清楚的，但是其确切叙述与数学假定都较为复杂. 我们在此只作一些提纲式的介绍. 读者完全可以跳过它们不读. 有兴趣的读者还可以查阅较高水平的随机过程教科书.

(1) 对于任意$\{\eta_n\}$停时 τ，**τ 前事件体**应该理解为，在 τ 时刻前"可以看到的"(掌握其信息的)随机事件全体. 将它记为$\mathscr{F}_\tau$，那么

$$\mathscr{F}_\tau \overset{\text{def}}{=} \{A: \text{对于任意 } n, \text{示性函数 } I_{A\cap\{\tau\leqslant n\}} \text{是} \boldsymbol{\eta}_n \text{ 可知的}\}.$$

(2) 如果$(\xi_n: n\geqslant 0)$是(η_n)鞅列，且存在 $p>1$ 使 $E|\xi_n|^p$ 对于 n 有界，则

$$\xi_{+\infty} \overset{\text{def}}{=} \lim_{n\to+\infty} \xi_n$$

概率为 1 地存在，且对于任意$\{\eta_n\}$停时 τ,σ，只要 $\tau>\sigma$，就有条件期望

$$E(\xi_\tau \mid \mathscr{F}_\sigma) = \xi_\sigma.$$

(3) 如果$(\xi_n: n\geqslant 0)$是(η_n)鞅列，则对于任意有界的$\{\eta_n\}$停时 τ,σ，只要 $\tau>\sigma$，就有条件期望

$$E(\xi_\tau \mid \mathscr{F}_\sigma) = \xi_\sigma.$$

(4) 如果$(\xi_n: n\geqslant 0)$是(η_n)鞅列，则对于任意$\{\eta_n\}$停时 τ，$(\xi_{\tau\wedge n}: n\geqslant 0)$是$(\eta_{\tau\wedge n})$鞅列. 且进一步可以得到更强的数学结论：$(\xi_{\tau\wedge n}: n\geqslant 0)$是$(\eta_n)$鞅列.

3.2.5 下鞅列的 Doob 分解

命题 3.5(下鞅列的 Doob 分解) 设$\{\xi_n\}$是(η_n)下鞅列，则存在唯一的(η_n)鞅列 M_n，满足：

(1) 初值 $M_1=\xi_1$.

(2) 对于$\boldsymbol{\eta}_n \overset{\text{def}}{=} \{\eta_1,\eta_2,\cdots,\eta_n\}$，存在$(\eta_{n-1})$可知的(也称为$(\eta_n)$**可料的**)递增随机序列 A_n，使 $A_1=0$，且

$$\xi_n = M_n + A_n.$$

证明 先证明存在性. 令

$$A_1 = 0, \quad A_n = \sum_{k=1}^{n-1}[E(\xi_{k+1} \mid \boldsymbol{\eta}_k) - \xi_k] \quad (n\geqslant 2)$$

由此定义得到随机序列$\{A_n\}$是(η_{n-1})可知的. 用下鞅性得到$\{A_n\}$是递增的. 再令

$$M_1 = \xi_1, \quad M_n = \xi_n - \sum_{k=1}^{n-1} [E(\xi_{k+1} \mid \boldsymbol{\eta}_k) - \xi_k] \quad (n \geqslant 2),$$

那么,容易验证它是鞅列.

再证明唯一性.设另有一个符合条件的分解 $\xi_n = M_n' + A_n'$,那么鞅列

$$M_n - M_n' = A_n' - A_n$$

应该是(η_{n-1})可知的.所以

$$M_n - M_n' = E[(M_n - M_n') \mid \boldsymbol{\eta}_{n-1}] = M_{n-1} - M_{n-1}' = \cdots = M_1 - M_1' = 0.$$

3.3　连续时间参数的鞅

3.3.1　连续时间参数的随机事件的历史参照与连续时间参数的鞅

定义 3.7　用一个随机过程$\{\eta_t: t \geqslant 0\}$作为整个"历史事件"的参照.令

$$\boldsymbol{\eta}_t \overset{\text{def}}{=\!=} \{\eta_s: s \leqslant t\}.$$

随机变量 ξ 称为(η_t)**可知的**,如果 ξ 可以表示成$\boldsymbol{\eta}_t$ 中有限个元素的函数列的极限,即 $\xi \in \bar{\Phi}(\boldsymbol{\eta}_t)$,其中 $\bar{\Phi}(\boldsymbol{\eta}_t)$是$\boldsymbol{\eta}_t$ 中任意有限个随机变量的任意函数列的极限全体.随机过程$\{\xi_t: t \geqslant 0\}$称为(η_t)**可知的**,或者(η_t)**适应的**,如果对于任意 t,随机变量 ξ_t 是(η_t)**可知的**.

定义 3.8　随机过程$\{\eta_t: t \geqslant 0\}$称为鞅,如果任意 t,η_t 的数学期望有限,且对于任意 n,任意 $t > s > s_1 > \cdots > s_n$ 有

$$E(\eta_t \mid \eta_s, \eta_{s_1}, \cdots, \eta_{s_n}) = \eta_s.$$

另一个数学期望函数有限的随机过程$\{\xi_t: t \geqslant 0\}$称为(η_t)鞅,如果$\{\xi_t\}$为(η_t)可知的,且对于任意 n,任意 $t > s > s_1 > \cdots > s_n$ 有

$$E(\xi_t \mid \eta_s, \eta_{s_1}, \cdots, \eta_{s_n}) = \xi_s.$$

由条件期望的性质,可以证明这个定义等价于:对于任意 $t > s$ 有

$$E(\xi_t \mid \eta_u: u \leqslant s) = \xi_s.$$

由条件期望的性质直接得到

若$\{\xi_t\}$是(η_t)鞅,则$\{\xi_t\}$也是鞅.

与鞅列类似地,连续时间的鞅的直观含义仍然是公平博弈.

定义 3.9　初始值 $\xi_0 = 0$ 的随机过程$\{\xi_t: t \geqslant 0\}$称为**独立增量过程**,如果对于任意 n,在任意互不相交的区间$(s_1, t_1]$,$(s_2, t_2]$,$\cdots$,$(s_n, t_n]$上的 $\xi_{t_1} - \xi_{s_1}, \xi_{t_2} - \xi_{s_2}, \cdots, \xi_{t_n} - \xi_{s_n}$ 都相互独立.

独立增量过程是随机徘徊在时间参数为连续情形时的推广.

例 3.10　数学期望为 0 的独立增量过程 η_t 是鞅.

证明　对于任意 n，任意 $t>s>s_1>\cdots>s_n$，$\eta_0=0$ 当然与 $\{\eta_s-\eta_{s_1},\eta_{s_1}-\eta_{s_2},\cdots,\eta_{s_n}-\eta_0\}$ 独立，而后者又与 $\{\eta_s,\eta_{s_1},\cdots,\eta_{s_n}\}$ 能相互地线性表出，所以 $\eta_t-\eta_s$ 就与 $\{\eta_s,\eta_{s_1},\cdots,\eta_{s_n}\}$ 独立. 利用条件期望的性质，有

$$
\begin{aligned}
E(\eta_t \mid \eta_s,\eta_{s_1},\cdots,\eta_{s_n}) &= E(\eta_t-\eta_s \mid \eta_s,\eta_{s_1},\cdots,\eta_{s_n})+E(\eta_s \mid \eta_s,\eta_{s_1},\cdots,\eta_{s_n}) \\
&= E(\eta_t-\eta_s)+\eta_s=\eta_s.
\end{aligned}
$$

作为例 3.10 的特例，我们得到下面一些讨论.

例 3.11　若 $\{N_t: t\geqslant 0\}$ 为强度为 λ 的 Poisson 过程，则 $\{N_t-\lambda t: t\geqslant 0\}$ 是鞅.

例 3.12(独立增量过程的平方可积鞅)　假定例 3.10 中的独立增量过程有二阶矩，即对于任意 t，$\sigma_t^2 \overset{\text{def}}{=} E\eta_t^2<+\infty$. 定义

$$
\xi_t=\eta_t^2-\sigma_t^2,
$$

那么 $\{\xi_t: t\geqslant 0\}$ 为 (η_t) 鞅.

证明　显然随机过程 $\{\xi_t: t\geqslant 0\}$ 为 (η_t) 可知的. 由于 $(\eta_t-\eta_s)^2$ 与 $(\eta_s,\eta_{s_1},\cdots,\eta_{s_n})$ 独立. 因而，对于任意 n，任意 $t>s>s_1>\cdots>s_n$，有

$$
\begin{aligned}
E(\eta_t^2-\eta_s^2 \mid \eta_s,\eta_{s_1},\cdots,\eta_{s_n}) &= E([(\eta_t-\eta_s)^2+2(\eta_t-\eta_s)\eta_s] \mid \eta_s,\eta_{s_1},\cdots,\eta_{s_n}) \\
&= E(\eta_t-\eta_s)^2+2\eta_s E(\eta_t-\eta_s)=E(\eta_t-\eta_s)^2 \\
&= E\eta_t^2-E\eta_s^2-2E[\eta_s(\eta_t-\eta_s)]=\sigma_t^2-\sigma_s^2.
\end{aligned}
$$

命题 3.6　设 $\{\xi_t: t\geqslant 0\}$ 为 (η_t) 鞅，则 $E\xi_t\equiv E\xi_0$.

定义 3.10　可以取值 $+\infty$ 的非负随机时刻 τ，称为**关于 $\{\eta_t\}$ 的停时**(简称 **$\{\eta_t\}$ 停时**)，如果对于任意 t，事件 $\{\tau$ 在 t 前发生与否$\}$(即事件 $\{\tau\leqslant t\}$)只依赖于 η_t 的历史记录，或者说，随机变量 $I_{\{\tau\leqslant t\}}$ 是 (η_t) 可知的. 在不会混淆的情形下，$\{\eta_t\}$ 停时常简称为**停时**.

显见，常数时间是停时；又如果 τ,σ 都是停时，$t\geqslant 0$，则 $\tau\wedge\sigma$，$\tau\wedge t$ 也是停时.

3.3.2　连续时间鞅的选样定理

与离散时间的鞅列类似地，对于连续时间的鞅，同样有如下的选样定理

定理 3.7(选样定理)　设 $\xi_t(t\geqslant 0)$ 是轨道连续的(其实只需轨道右连续)鞅(即对于任意基本事件 ω，样本轨道 $\xi_t(\omega)$ 是 t 的连续函数). 又若停时 τ 满足：

(1) τ 是有界停时.

或更一般地

(2) ξ_τ 有有限的期望，且 $\lim\limits_{t\to+\infty} E(|\xi_t|\cdot I_{\{\tau>t\}})=0$.

那么，有

$$
E\xi_\tau=E\xi_0.
$$

同样，也有更广的 Doob 停止定理与下鞅分解定理，不过后者需要加上一个较为复杂

的数学条件，而且定理的叙述也复杂得多，称为 Doob-Meyer 分解. 它是现代鞅论的起点.

定理 3.8　如果对于任意 n 固定，$\{\xi_t^{(n)}\}$ 都是 (η_t) 鞅. 又如果对于任意 t 都有

$$E\mid \xi_t^{(n)}-\xi_t\mid \to 0,$$

那么，ξ_t 也是 (η_t) 鞅.

证明　由全期望公式，对于 $s<t$，有

$$E\mid E(\xi_t^{(n)}\mid \eta_u:u\leqslant s)-E(\xi_t\mid \eta_u:u\leqslant s)\mid$$

$$\leqslant E[E(\mid \xi_t^{(n)}-\xi_t\parallel \eta_u:u\leqslant s)]=E\mid \xi_t^{(n)}-\xi_t\mid \to 0.$$

由假定 $\{\xi_t^{(n)}\}$ 为 (η_t) 鞅，即对 $s<t$，有

$$\xi_s^{(n)}=E(\xi_t^{(n)}\mid \eta_u:u\leqslant s),$$

令 $n\to +\infty$，便得

$$E(\xi_t\mid \eta_u:u\leqslant s)=\xi_s.$$

3.3.3　平方可积鞅

定义 3.11　$\{\xi_t\}$ 称为 (η_t) 平方可积鞅，如果它是 (η_t) 鞅，且对于任意 t 固定，ξ_t 的方差有限.

引理 3.9

(1) 如果 ξ,η 方差有限，则有 Schwartz 不等式

$$\mid E(\xi\eta)\mid^2\leqslant E\xi^2E\eta^2.$$

(2) 如果随机变量列 $\{\xi_n\}$，$\{\eta_n\}$ 满足

$$E(\xi_n-\xi)^2\to 0,\quad E(\eta_n-\eta)^2\to 0,$$

那么

$$E(\xi_n\eta_n)\to E(\xi\eta).$$

(3) 如果 ξ 方差有限，则条件期望有 Schwartz 不等式

$$\mid E(\xi\mid \eta_u:u\leqslant s)\mid^2\leqslant E(\xi^2\mid \eta_u:u\leqslant s).$$

(4) 如果随机变量列 $\{\xi_n\}$ 满足 $E(\xi_n-\xi)^2\to 0$，那么

$$E[E(\xi_n\mid \eta_u:u\leqslant s)-E(\xi\mid \eta_u:u\leqslant s)]^2\to 0.$$

证明　(1) $E(\xi+t\eta)^2$ 是 t 的非负二次多项式，其判别式

$$4\mid E(\xi\eta)\mid^2-4E\xi^2E\eta^2\leqslant 0.$$

(2) 由

$$E\eta_n^2\leqslant 2[E(\eta_n-\eta)^2+E\eta^2]$$

对于 n 是有界的，利用(1)得到

$$\mid E(\xi_n\eta_n)-E(\xi\eta)\mid=\mid E[(\xi_n-\xi)\eta_n]+E[\xi(\eta_n-\eta)]\mid$$

$$\leqslant E\mid(\xi_n-\xi)\eta_n\mid+E\mid\xi(\eta_n-\eta)\mid$$

$$\leqslant \sqrt{E(\xi_n-\xi)^2}\sqrt{E\eta_n^2}+\sqrt{E\xi^2}\sqrt{E(\eta_n-\eta)^2}\to 0.$$

(3) 仿照(1)便得.

(4) 由(3)及全期望公式推出

$$\begin{aligned}&E[E(\xi_t^{(n)}\mid\eta_u: u\leqslant s)-E(\xi_t\mid\eta_u: u\leqslant s)]^2\\&\quad=E[E(|\xi_t^{(n)}-\xi_t\|\eta_u: u\leqslant s)^2]\leqslant E[E([\xi_t^{(n)}-\xi_t]^2\mid\eta_u: u\leqslant s)]\\&\quad=E(\xi_t^{(n)}-\xi_t)^2\to 0.\end{aligned}$$

推论 3.10 如果$\{\xi_t^{(n)}\}$是(η_t)平方可积鞅,且对于任意 t,$E(\xi_t^{(n)}-\xi_t)^2\to 0$,那么,$\{\xi_t\}$也是$(\eta_t)$平方可积鞅,而且

$$E[E(\xi_n\mid\eta_u: u\leqslant s)-E(\xi\mid\eta_u: u\leqslant s)]^2\to 0.$$

证明 由于

$$E\mid\xi_t^{(n)}-\xi_t\mid\leqslant\sqrt{E\mid\xi_t^{(n)}-\xi_t\mid^2}\to 0,$$

故定理 3.8 的假定成立,所以$\{\xi_t\}$是(η_t)鞅. 又由假定 $E(\xi_s^{(n)}-\xi_s)^2\to 0$ 可知 $E\xi_t^2<+\infty$.

定理 3.11(连续鞅的 Kolmogorov 极值不等式) 设$\{\xi_t, t\geqslant 0\}$是轨道连续的鞅,那么对于任意 $\varepsilon>0$ 成立不等式

$$P(\max_{0\leqslant t\leqslant T}\xi_t\geqslant\varepsilon)\leqslant\frac{E(\max\{0,\xi_T\})}{\varepsilon}.$$

证明 对于任意 n,以及任意 $t_1,t_2,\cdots,t_n\leqslant T$,事件组

$$B_i=\{\xi_{t_k}<\varepsilon, k<i, \xi_{t_i}\geqslant\varepsilon\}$$

是事件

$$B\stackrel{\text{def}}{=}\{\max_{0\leqslant t\leqslant T}\xi_t>\varepsilon\}$$

的一组划分. 我们有

$$\begin{aligned}E(\max\{0,\xi_T\})&\geqslant E(\max\{0,\xi_T\}I_B)=\sum_{i=1}^n E[\max\{0,\xi_T\}I_{B_i}]\\&\geqslant\sum_{i=1}^n E[\xi_T I_{B_i}]=\sum_{i=1}^n E[I_{B_i}E(\xi_T\mid\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_i})]=\sum_{i=1}^n E[I_{B_i}\xi_{t_i}]\\&\geqslant\sum_{i=1}^n\varepsilon P(B_i)=\varepsilon P\{\max_{0\leqslant t\leqslant T}\xi_t>\varepsilon\}.\end{aligned}$$

让分点充分多,再取极限就得到定理.

注 仔细检查此证明可以发现,此不等式对于连续轨道的下鞅仍然成立,并称为下鞅的 Doob 不等式.

定理 3.11 的离散时间形式为如下的定理.

定理 3.12(鞅列的 Kolmogorov 不等式) 设$\{\xi_n, n\geqslant 1\}$是鞅列,则对任意 $\varepsilon>0$ 有

$$P(\max_{1\leqslant n\leqslant N}\xi_n\geqslant\varepsilon)\leqslant\frac{E[\max\{0,\xi_N\}]}{\varepsilon}.$$

实际上，建立鞅列的 Kolmogorov 不等式的构思，源于它的特殊情形，即下面的推论.

推论 3.13(独立和列的 Kolmogorov 不等式) 设$\{\xi_n,n\geqslant 1\}$是方差有限、均值为 0 的独立同分布随机序列的部分和序列

$$\xi_n=\sum_{k=1}^{n}\eta_k,$$

那么对于任意 $\varepsilon>0$ 有

$$P(\max_{1\leqslant n\leqslant N}\xi_n\geqslant\varepsilon)\leqslant\frac{E\xi_N^2}{\varepsilon^2}.$$

习题 3

1. 若$\{\xi_n\}$是鞅列，则 $\xi_n\vee c$ 是下鞅列，$\xi_n\wedge c$ 是上鞅列.

2. 若$\{\xi_n\}$是鞅列，而且 $E\xi_n^+<+\infty$，则$\{\xi_n^+\}$也是下鞅列. 对应地，又若 $E\xi_n^->-\infty$，则$\{\xi_n^-\}$也是上鞅列.

3. 设 ξ_t,η_t 是鞅列，则 $\xi_t+\eta_t$ 是鞅列，$\min\{\xi_t,\eta_t\}$是下鞅列.

4. 设 $\xi_n=\sum_{k=1}^{n}X_k$ 是鞅列，且其方差有限，则序列$\{X_k\}$中的随机变量两两不相关.

5. 在博采输光问题中，如果 $p<q$，求输光概率 p_a,q_b 以及博采至输光为止的平均时间.

6. 设$\{X_n,n\geqslant 1\}$独立同分布，$X_n\sim\begin{pmatrix}-1 & 1\\ q & p\end{pmatrix}$，$p>q$. $\xi_n=\sum_{k=1}^{n}X_k$，

$$\eta_n=\xi_n^2-n\left[1-\frac{n-1}{2}(p-q)^2\right],\quad \zeta_n=\left(\frac{q}{p}\right)^{\xi_n},\quad Z_n=\xi_n-n(p-q).$$

(1) 证明$\{\xi_n\}$是(X_n)下鞅列，而$\{\eta_n\}$，$\{\zeta_n\}$，$\{Z_n\}$都是(ξ_n)鞅列.

(2) 求 η_m 与 η_n 的相关系数.

7. 若 $X_n\sim N(\mu,\sigma^2)$，$\xi_n=\exp\left[-\frac{\mu}{\sigma^2}\left(\sum_{k=1}^{n}X_k-\frac{n\mu^2}{2\sigma^2}\right)\right]$. 证明$\{\xi_n\}$是$(X_n)$鞅列.

8. 假定随机序列$\{\xi_n,n\geqslant 0\}$有数学期望，且满足

$$E(\xi_{n+1}\mid\xi_n,\cdots,\xi_0)=\alpha\xi_n+\beta\xi_{n-1}\quad(\alpha,\beta>0,\alpha+\beta=1).$$

能否选取 c，使 $\eta_n=c\xi_n+\xi_{n-1}(n\geqslant 1,\eta_0=\xi_0)$是$(\xi_n)$鞅列?

9. 设 $\xi_0\sim U[0,1]$，又在 ξ_n 已知的条件下，$\xi_{n+1}\sim U[1-\xi_n,1]$，$\eta_0=\xi_0$，$\eta_n=2^n\prod_{k=1}^{n}\frac{1-\xi_k}{\xi_{k-1}}$. 证明$\{\eta_n\}$是$(\xi_n)$鞅列.

10. 设 $S_n=X_1+X_2+\cdots+X_n$ 是简单随机徘徊. $X_n\sim\begin{pmatrix}-1 & 1\\ q & p\end{pmatrix}(q=1-p)$且独立同分布. 证明 $\xi_n=(pz^{2q}+qz^{-2p})^{-n}z^{S_n-n(p-q)}$ 是鞅列.

11. 若$\{\xi_1,\xi_2,\cdots,\xi_n,\cdots\}$独立同分布,而正整值随机变量 η 与它们独立,那么有

$$E(\xi_1+\xi_2+\cdots+\xi_\eta)=E\xi_1E\eta,$$

$$\operatorname{var}(\xi_1+\xi_2+\cdots+\xi_\eta)=E\eta\operatorname{var}\xi_1+\operatorname{var}(\eta)(E\xi_1)^2.$$

第 4 章　Brown 运动与 Markov 过程

4.1　Brown 运动的数学模型

英国生物学家 Brown 于 1827 年在显微镜下观察液体中的花粉微粒，发现它们在极端不规则地运动. 以后的研究者发现了更多的类似现象，如空气中的烟雾的扩散等. 直至 19 世纪末才知道其机理是，由于花粉、烟尘等微粒，受到大量液体分子或气体分子的作用，所作的无规则碰撞而形成的. Einstein 在 1905 年作了量化的讨论，建立了物理模型. 以后又经过 Ornstein 和 Uhlenbeck，以及 Langevin 等人的完善. Wiener 在 1918 年对 Brown 运动建立了用随机过程的语言描述的严格数学模型. 此后，Brown 运动这个术语就专指 Wiener 的数学模型. 因此，Brown 运动也称为 Wiener 过程. 而早在 1900 年 Bachellier 在研究法国的债券市场的规律时，曾提出了一个相当于 Brown 现象的直观模型. 不幸的是这个先驱的模型在当时被忽略了.

经过数学家一个世纪的努力，到 20 世纪中叶，Brown 运动的数学理论已经十分成熟而且具有极为广泛的应用. Brown 运动和 Poisson 过程自然地成为随机过程的两大基柱.

Brown 运动在数值计算中，则可以用离散的对称简单随机徘徊近似.

作为随机过程，Brown 运动的性质最为特殊，作用也更为广泛. 在理论领域和应用领域中，Brown 运动的不仅出现在概率论领域，还遍及数学、物理、化学、生物、地球、天文等自然科学领域，数量经济、金融、精算等应用领域，以及人文科学的各个领域的计算与建模中.

Einstein 将 Brown 发现的现象描述为一个作随机运动的粒子在时间区间$[0,t]$上的随机位移，因此，它是一组依赖于时间参数 t 的三维随机向量$\{B_s: 0\leqslant s\leqslant t\}$，即它是一个三维的随机过程. 如果我们将粒子的出发位置取为坐标原点，那么 $B_0=0$. Einstein 从物理的角度假定了这种粒子的运动具有下性质：

（1）粒子在空间的位移的 3 个一维分量是相互独立的（我们可以只考虑其一个坐标分量. 为符号的简便. 我们仍将此分量记为$\{B_t, t\geqslant 0\}$），且它是独立增量过程，即在任意互不相交的区间$(s_1, t_1]$，$(s_2, t_2]$，…，$(s_n, t_n]$上，其差 $B_{t_1}-B_{s_1}$，$B_{t_2}-B_{s_2}$，…，$B_{t_n}-B_{s_n}$ 都相互独立.

（2）运动的统计规律对空间是对称的. 因而有 $EB_t=0$.

(3) 在时间区间上的差 $B_{t+s}-B_s$ 的分布，与时间区间的起点 s 无关，并且其方差

$$\sigma(t) \stackrel{\text{def}}{=} E\,(B_{t+h}-B_h)^2$$

是 t 的连续函数.

由此我们求 B_t 的分布密度的表达式如下.

首先，由独立增量性质得到

$$\begin{aligned}\sigma(t+s) &= E(B_{t+s}-B_0)^2 = E(B_{t+s}-B_t+B_t-B_0)^2 \\ &= E(B_{t+s}-B_t)^2 + E(B_t-B_0)^2 = \sigma(s)+\sigma(t).\end{aligned}$$

由于 $\sigma(t)$ 关于 t 连续，由微积分知道 $\sigma(t)$ 必有表达式

$$\sigma(t) = Dt,$$

其中 D 是一个常数，是单位时间内粒子平方位移的均值，称之为扩散常数. 由分子运动学 Einstein 得到了

$$D = \frac{RT}{Nf},$$

其中 R 是由分子的特性所决定的一个普适常数，T 是绝对温度，N 是 Avogadro 常数，f 是摩擦系数. 在以下讨论中，我们不妨取 $D=1$.

对任意的划分

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t,$$

B_t 可表示为 n 个独立的随机变量之和：

$$B_t = (B_{t_n}-B_{t_{n-1}}) + (B_{t_{n-1}}-B_{t_{n-2}}) + \cdots + (B_{t_1}-B_{t_0}).$$

命题 4.1 B_t 具有正态密度 $N(0,t)$，即其分布密度为

$$p(t,x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}.$$

证明 我们用特征函数来推导 B_t 的分布密度 $p(t,x)$. 记 B_t 的特征函数为

$$\varphi(t,\theta) = Ee^{i\theta B_t} \quad (-\infty<\theta<+\infty).$$

利用独立增量性质，得到

$$\begin{aligned}\varphi(t+s,\theta)-\varphi(t,\theta) &= Ee^{i\theta B_{t+s}} - Ee^{i\theta B_t} \\ &= E[e^{i\theta B_t}(e^{i\theta(B_{t+s}-B_t)}-1)] = E[e^{i\theta(B_t-B_0)}(e^{i\theta(B_{t+s}-B_t)}-1)] \\ &= E(e^{i\theta(B_t-B_0)})E(e^{i\theta(B_{t+s}-B_t)}-1) = Ee^{i\theta B_t}E(e^{i\theta(B_s-B_0)}-1) \\ &= \varphi(t,\theta)E(e^{i\theta B_s}-1).\end{aligned} \tag{4.1}$$

直观地，在 Taylor 展开式(严格的推导还需要假定$\lim\limits_{s\to 0}E|B_s|^3=0$)

$$e^{ix} = 1+ix-\frac{x^2}{2}+o(x^2) \quad (x\to 0) \tag{4.2}$$

中令 $x=\theta B_s$，再取期望，由(4.2)式得到

$$E(e^{i\theta B_s}-1) = i\theta EB_s - \frac{\theta^2}{2}EB_s^2 + o(EB_s^2)$$

$$=-\frac{1}{2}\theta^2\sigma(s)+o(s).$$

对(4.1)式等号的两边都除以 s 后，令 $s\to 0$，便得

$$\frac{\partial\varphi}{\partial t}(t,\theta)=-\frac{1}{2}\theta^2\varphi(t,\theta).$$

这是关于 t 的一个一阶的常系数常微分方程，在初始条件

$$\varphi(0,\theta)=E(\mathrm{e}^{\mathrm{i}\theta B_0})=1$$

下，求解这个方程便得

$$\varphi(t,\theta)=\mathrm{e}^{-\frac{1}{2}\theta^2 t}.$$

它恰是正态分布 $N(0,t)$的特征函数. 从而由特征函数与分布的一一对应性可知，B_t 的分布密度为 $N(0,t)$.

从 Einstein 的物理模型可以抽象出以下的数学模型.

定义 4.1　满足以下条件的一个随机过程$\{B_t,t\geqslant 0\}$称为 **Brown 运动**：

(1) $B_0=0$，$\{B_t,t\geqslant 0\}$是独立增量过程，即对任意互不相交的区间$(s_1,t_1]$，$(s_2,t_2]$，…，$(s_n,t_n]$，相应的增量 $B_{t_1}-B_{s_1}$，$B_{t_2}-B_{s_2}$，…，$B_{t_n}-B_{s_n}$ 都相互独立；

(2) 对于任意 $s\geqslant 0$，$t>0$，增量 $B_{s+t}-B_s\sim N(0,Dt)$(分布不依赖 s，称为具有**平稳增量**)；

(3) 对每一个固定的基本事件(样本点)ω，$B_t(\omega)$作为 t 的函数(称为样本轨道)，是连续函数(微粒运动的连续性).

特别地，当 $D=1$ 时，我们称之为**标准 Brown 运动**. 一般就简称标准 Brown 运动为 Brown 运动.

注　有时 B_t，$B_t(\omega)$也分别用 $B(t)$，或 $B(t,\omega)$来表示，称为 Brown 运动的轨道，Brown 运动的典型轨道为图 4.1 所示.

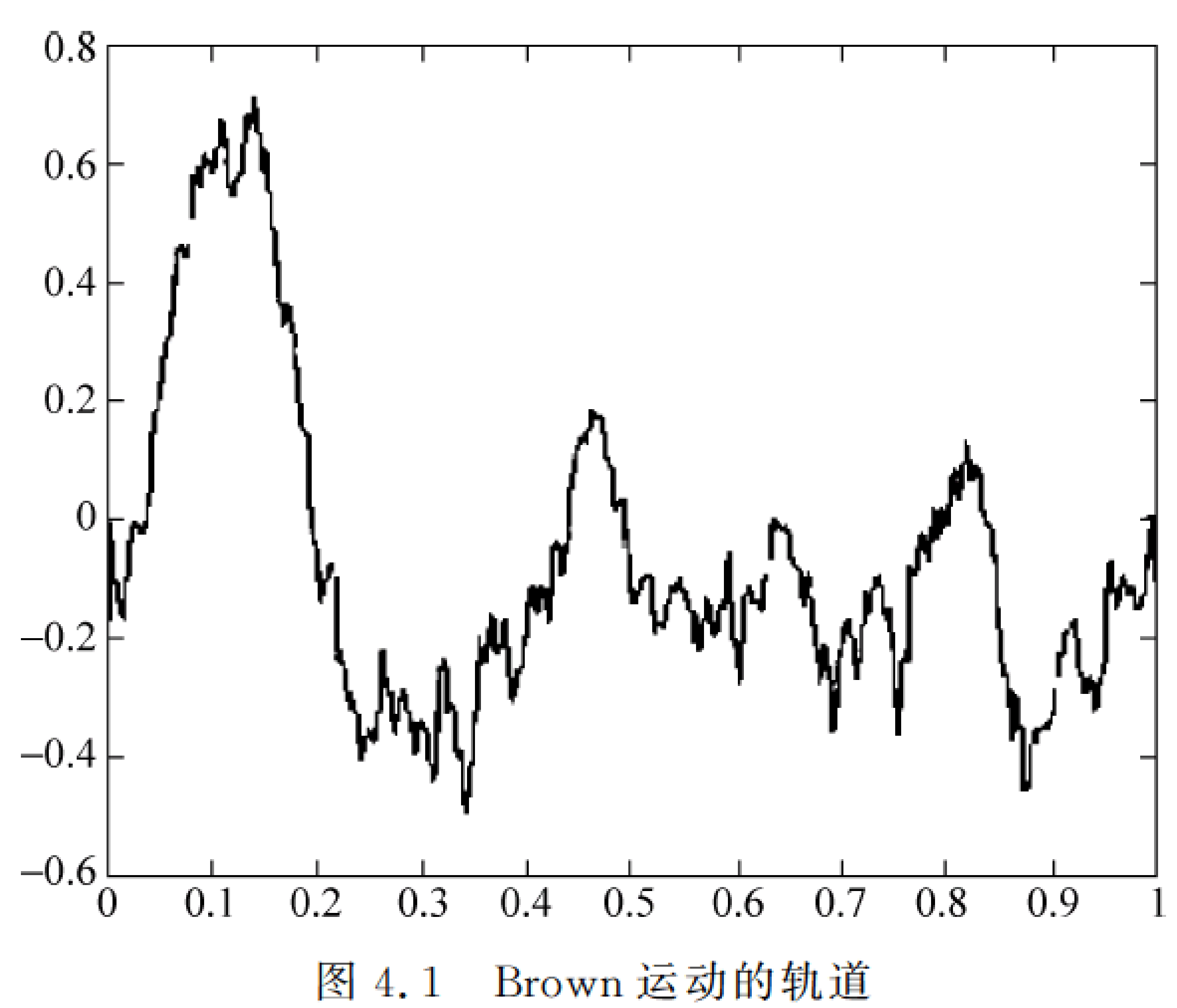


图 4.1　Brown 运动的轨道

注　Brown 运动中的条件(3)的合理性(即加上它不会引起矛盾),可以通过概率论理论证明.在本书中我们将承认它而并不给出其证明,而在实用中它常常是被认为当然成立而忽视其证明的.

命题 4.2　对于 Brown 运动 B_t,有

(1) B_t 是鞅;

(2) (平方可积鞅)B_t^2-t 是(B_t)鞅;

(3) (指数鞅)$z_t \stackrel{\text{def}}{=\!=} e^{cB_t-\frac{c^2t}{2}}$是$(B_t)$鞅.

证明　(1)和(2)是第 3 章的特例.今验证(3).对于任意 n,任意 $t>s>s_1>\cdots>s_n$,由 Brown 运动的独立增量性质知道$\dfrac{z_t}{z_s}=e^{c(B_t-B_s)-\frac{c^2(t-s)}{2}}$与$(B_s,B_{s_1},\cdots,B_{s_n})$独立.又因为 z_s 是 B_s 的函数,利用条件期望的性质便得

$$
\begin{aligned}
E(z_t\mid B_s,B_{s_1},\cdots,B_{s_n})&=z_sE\left(\frac{z_t}{z_s}\middle|B_s,B_{s_1},\cdots,B_{s_n}\right)\\
&=z_sE\left(e^{c(B_t-B_s)-\frac{c^2(t-s)}{2}}\mid B_s,B_{s_1},\cdots,B_{s_n}\right)\\
&=z_sE\left(e^{c(B_t-B_s)-\frac{c^2(t-s)}{2}}\right)=z_s.
\end{aligned}
$$

命题 4.3　独立增量过程$\{\xi_t: t\geqslant 0\}$具有以下的特性:过程的增量 $\xi_{t+s}-\xi_s$ 与过程过去的信息$\{\xi_u: 0\leqslant u<s\}$是独立的,其含义为:对于任意 t,过程的增量 $\xi_{t+s}-\xi_s$ 与 $\bar{\Phi}(\xi_t)$中的任意随机变量独立.进而有

$$P(\xi_{t+s}\leqslant y\mid \xi_s=x,\xi_u=x_u(0\leqslant u<s))=P(\xi_{t+s}\leqslant y\mid \xi_s=x).$$

证明　利用独立增量性质,有

$$
\begin{aligned}
&P(\xi_{t+s}\leqslant y\mid \xi_s=x,\xi_u=x_u(0\leqslant u<s))\\
&\quad=P(\xi_{t+s}-\xi_s\leqslant y-x\mid \xi_s=x,\xi_u=x_u(0\leqslant u<s))\\
&\quad=P(\xi_{t+s}-\xi_s\leqslant y-x)=P(\xi_{t+s}\leqslant y\mid \xi_s=x).
\end{aligned}
$$

4.2　Markov 过程与 Brown 运动的 Markov 性

定义 4.2　随机过程$\{\xi_t,t\geqslant 0\}$称为 **Markov 过程**,如果对于任意的 $t,s,y,x,x_u(0\leqslant u<s)$,满足

$$P(\xi_{t+s}\leqslant y\mid \xi_s=x,\xi_u=x_u(0\leqslant u<s))=P(\xi_{t+s}\leqslant y\mid \xi_s=x).$$

其直观含义是:在已知随机过程现在的情形 $\xi_s=x$ 的条件下,过程的过去情形 $\xi_u=x_u(0\leqslant u<s)$与随机过程的将来情形 $\xi_{t+s}=y$ 是相互条件独立的.如果 Markov 过程还满足条件

$$P(\xi_{t+s}\leqslant y\mid \xi_s=x)=P(\xi_t\leqslant y\mid \xi_0=x),$$

则称为**时齐的 Markov 过程**.

定义 4.3 假定$\{\xi_t, t\geqslant 0\}$是 Markov 过程. 我们称

$$F(s,x;\ t,y) \stackrel{\text{def}}{=\!=} P(\xi_t \leqslant y \mid \xi_s = x)$$

为$\{\xi_t, t\geqslant 0\}$的**转移分布函数**. 而将 $F(s,x;\ t,y)=P(\xi_t\leqslant y|\xi_s=x)$对于 y 的密度函数 $p(s,x;\ t,y)$(如果存在),称为$\{\xi_t, t\geqslant 0\}$的**转移密度函数**,简称**转移密度**. 对于时齐的 Markov 过程,我们简记

$$F(t,x,y) = F(s,x;\ s+t,y),$$
$$p(t,x,y) = p(s,x;\ s+t,y)$$

由命题 4.3 可知,独立增量过程(即具有独立增量性质的随机过程)是 Markov 过程.

命题 4.4 如果 Markov 过程存在转移密度,那么它满足:

(P.1) 对于任意 $s<t$ 有

$$p(s,x;\ t,y) \geqslant 0, \quad \int p(s,x;\ t,y)\mathrm{d}y = 1;$$

(P.2) Chapman-Kolmogorov 方程:对于 $s<u<t$

$$p(s,x;\ t,y) = \int p(s,x;\ u,z)p(u,z;\ t,y)\mathrm{d}z.$$

而对时齐的 Markov 过程,就变成

(P.1) 对于任意 t 有

$$p(t,x,y) \geqslant 0, \quad \int p(t,x,y)\mathrm{d}y = 1;$$

(P.2) Chapman-Kolmogorov 方程:对于任意 s,t 有

$$p(t+s,x,y) = \int p(s,x,z)p(t,z,y)\mathrm{d}z.$$

这些公式的证明需要利用条件期望的技巧,这些技巧并不困难,但是因为这里的等式的含义非常直观,故而我们略去其证明.

一个随机过程在任意有限个时刻上的联合分布全体,称为此随机过程的**有限维分布族**.

命题 4.5 如果 Markov 过程$\{\xi_t, t\geqslant 0\}$具有转移密度 $p(t,x,y)$,且其初始随机变量 ξ_0 具有分布密度 $\varphi(x)$,则过程$\{\xi_t, t\geqslant 0\}$的一切有限维联合分布都有密度,即对于任意 n 及任意 $0<t_1<\cdots<t_n$,$(\xi_0,\xi_{t_1},\cdots,\xi_{t_n})$有联合密度.

$$f_{(\xi_0,\xi_{t_1},\cdots,\xi_{t_n})}(x_0,x_1,\cdots,x_n)$$
$$=\varphi(x_0)p(t_1,x_0,x_1)p(t_2-t_1,x_1,x_2)\cdots p(t_n-t_{n-1},x_{n-1},x_n).$$

又在 ξ_0 取定值 x 的条件下,$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$有联合密度

$$f_{(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})}(x_1,x_2,\cdots,x_n) = p(t_1,x,x_1)p(t_2-t_1,x_1,x_2)\cdots p(t_n-t_{n-1},x_{n-1},x_n).$$

证明 只需利用条件密度的乘积公式,以及条件密度的 Markov 性等式

$$p_{\xi_{t_{k+1}}}(x_{k+1} \mid \xi_{t_k} = x_k,\cdots,\xi_{t_1} = x_1,\xi_0 = x_0) = p_{\xi_{t_{k+1}}}(x_{k+1} \mid \xi_{t_k} = x_k).$$

推论 4.6 如果放弃 Brown 运动初值为 0 的假定，即 B_0 可以是随机变量，但是与 $\{B_t,t>0\}$ 独立，那么这样得到的 $\{B_t,t\geqslant 0\}$ 仍旧是时齐的 Markov 过程.

将 $\{B_t,t\geqslant 0\}$ 的转移密度记为 $b(t,x,y)$，其表达式为

$$b(t,x,y) \stackrel{\text{def}}{=\!=} p(t,y-x) \stackrel{\text{def}}{=\!=} \frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{(y-x)^2}{2t}}.$$

它满足以下的关系

$$b(t+s,x,y) = \int b(t,x,u)b(s,u,y)\mathrm{d}u.$$

这个关系实际上是转移密度的全概率公式，就是 Chapman-Kolmogorov 方程.

证明 在 $B_0=x$ 的条件下，利用 $B_t-x\sim N(0,t)$，便得到转移密度的表达式. 而 Chapman-Kolmogorov 方程可以通过积分演算直接得到.

4.3 Brown 运动的有限维联合密度与基本性质

命题 4.7 对任意的 $0=t_0<t_1<\cdots<t_{n-1}<t_n$，$(B_{t_1},B_{t_2},\cdots,B_{t_n})$ 在点 $(x_1,x_2,\cdots,x_n)$ 的联合密度为

$$f_{(B_{t_1},B_{t_2},\cdots,B_{t_n})}(x_1,x_2,\cdots,x_n) = \frac{\exp\left[-\left(\frac{x_1^2}{2t_1}+\frac{(x_2-x_1)^2}{2(t_2-t_1)}+\cdots+\frac{(x_n-x_{n-1})^2}{2(t_n-t_{n-1})}\right)\right]}{(2\pi)^{\frac{n}{2}}\sqrt{t_1(t_2-t_1)\cdots(t_n-t_{n-1})}}.$$

证明 由于 $B_0=0$，$B_t\sim N(0,t)$，对任意的 $0=t_0<t_1<\cdots<t_{n-1}<t_n$，由独立增量性知道，$(B_{t_1},B_{t_2},\cdots,B_{t_n})$ 在点 $(x_1,x_2,\cdots,x_n)$ 的联合密度为

$$\begin{aligned} f_{(B_{t_1},B_{t_2},\cdots,B_{t_n})}(x_1,x_2,\cdots,x_n) &= p(t_1,x_1)p(t_2-t_1,x_2-x_1)\cdots p(t_n-t_{n-1},x_n-x_{n-1}) \\ &= \frac{\exp\left[-\left(\frac{x_1^2}{2t_1}+\frac{(x_2-x_1)^2}{2(t_2-t_1)}+\cdots+\frac{(x_n-x_{n-1})^2}{2(t_n-t_{n-1})}\right)\right]}{(2\pi)^{\frac{n}{2}}\sqrt{t_1(t_2-t_1)\cdots(t_n-t_{n-1})}}. \end{aligned}$$

注 这个联合密度的主干部分是指数 $(x_1,x_2,\cdots,x_n)$ 的二次型，所以 $(B_{t_1},B_{t_2},\cdots,B_{t_n})$ 的联合分布是多维正态分布.

推论 4.8 Brown 运动是 Gauss 过程.

命题 4.9 Brown 运动 $\{B_t,t\geqslant 0\}$ 对于 $t_1<t_2$，有条件分布

$$B_{t_2}\,|_{B_{t_1}} \sim N(B_{t_1},t_2-t_1),$$

以及条件分布

$$B_{t_1}\,|_{B_{t_2}} \sim N\left(\frac{t_1}{t_2}B_{t_2},\frac{t_1}{t_2}(t_2-t_1)\right).$$

证明 第一个条件分布较为显然. 我们证明第二个条件分布.

对于 $s<t$，在条件 $B_t=y$ 下，B_s 的条件分布密度，记为 $f_{s|t}(x|y)$，其表达式为

$$f_{s|t}(x \mid y)=\frac{p(s,x)p(t-s,y-x)}{p(t,y)} \stackrel{\text{def}}{=} C_1\exp\left[-\frac{x^2}{2s}-\frac{(y-x)^2}{2(t-s)}\right]$$

$$\stackrel{\text{def}}{=} C_2(y)\exp\left[-x^2\left(\frac{1}{2s}+\frac{1}{2(t-s)}\right)+\frac{yx}{t-s}\right]$$

$$= C_2(y)\exp\left[-\frac{t}{2s(t-s)}\left(x^2-2\,\frac{sy}{t}x\right)\right]$$

$$\stackrel{\text{def}}{=} C_3(y)\exp\left[-\frac{1}{2s(t-s)/t}\left(x-\frac{sy}{t}\right)^2\right],$$

其中 C_1 为常数. 这正说明，在条件 $B_t=y$ 下，B_s 的条件分布是均值为 $\frac{s}{t}y$，方差为 $\frac{s}{t}(t-s)$ 正态分布：$N\left(\frac{s}{t}y,\frac{s}{t}(t-s)\right)$.

Brown 运动是一个 Gauss 过程，其分布特性由其均值函数和协方差函数完全决定. Brown 运动的均值函数是零函数，而利用其独立增量性可知其协方差函数为：当 $s\leqslant t$ 时有

$$C(s,t)=\operatorname{cov}(B_s,B_t)=\operatorname{cov}(B_s,B_s+B_t-B_s)$$
$$=\operatorname{cov}(B_s,B_s)+\operatorname{cov}(B_s,B_t-B_s)=R(s,s)=s.$$

可见，一般地有下面的命题.

命题 4.10　对 Brown 运动 $\{B_t,t\geqslant 0\}$ 有

$$\mu(t)=0,$$
$$C(s,t)=\operatorname{cov}(B_s,B_t)=E(B_sB_t)=s\wedge t.$$

反之，如果 Gauss 过程满足 $\mu(t)=0, C(s,t)=s\wedge t$，则它就是 Brown 运动.

证明　满足 $\mu(t)=0, C(s,t)=s\wedge t$ 的 Gauss 过程的有限维分布，与 Brown 运动的有限维分布是一样的，而轨道的连续性，则可以通过更为深入的随机过程的理论分析，可以通过合理的修正得到.

命题 4.11　对 Brown 运动 $\{B_t,t\geqslant 0\}$ 有

$$E(B_t-B_s)^2=|t-s|.$$

证明　不妨设 $s\leqslant t$. 由命题 4.7 得

$$E(B_t-B_s)^2=EB_t^2-2E(B_sB_t)+EB_s^2=t-2s+s=t-s,$$

故而结论成立.

从定义容易验证 Brown 运动有如下的事实.

命题 4.12　对 Brown 运动 $\{B_t,t\geqslant 0\}$ 有

(1) **平移的统计不变性**：对于常数 $s>0$，$\{B_t-B_s\}$ 在 $t\geqslant s$ 是 Brown 运动. 可参见图 4.2.

(2) **尺度的统计不变性**：对于常数 $c>0$，$\left\{\frac{B_{ct}}{\sqrt{c}},t\geqslant 0\right\}$ 也是 Brown 运动.

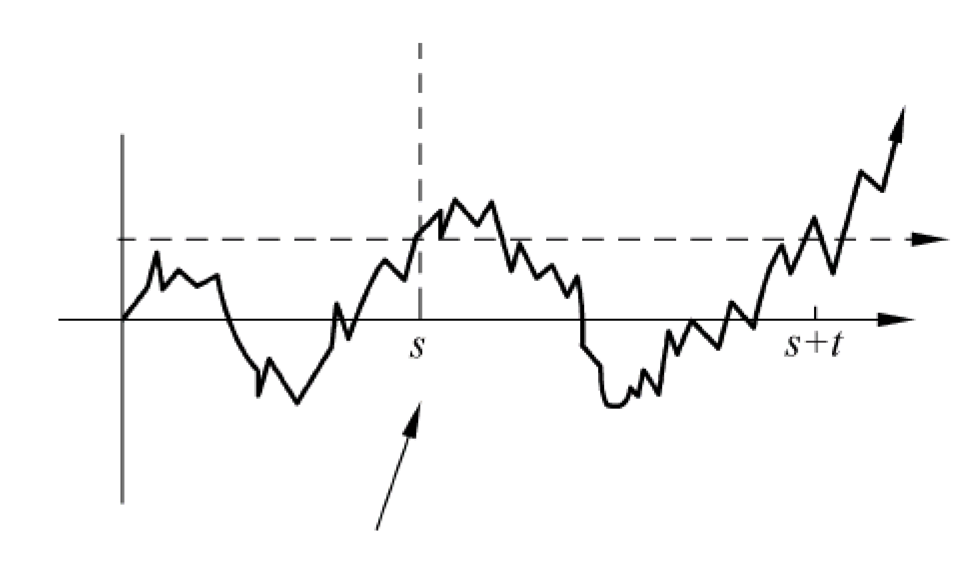


图 4.2　Brown 运动在时刻 s 重新开始

(3) $\{tB_{\frac{1}{t}}:t>0\}$ 也是 Brown 运动.

证明　只需验证它们都是 Gauss 过程，且满足命题 4.10 的条件.

Brown 运动有许多奇异的轨道性质. 由 Brown 运动 $\{B_t, t\geqslant 0\}$ 的定义，要求 $\{B_t, t\geqslant 0\}$ 的样本轨道(我们提醒：对每一个固定的样本点 ω，$B_t(\omega)$ 作为 t 的函数，称为对应于样本点 ω 的样本轨道)是连续函数. 然而，它的轨道却不是我们通常所见到的连续函数，而是一个处处都不可微的连续函数(它有明确的物理含义，即由于粒子在每一瞬间，都会受到介质中的分子的碰撞，碰撞后的粒子立即改变运动方向，因而速度为无穷，故而样本轨道是不可微的). 从数学模型分析，由 Brown 运动的定义，$B_{t_0+h}-B_{t_0}\sim N(0,h)$，即 $\dfrac{B_{t_0}-B_{t_0+h}}{\sqrt{h}}\sim N(0,1)$，故而对任意给定的正数 C，有

$$P\left(\omega:\left|\frac{B_{t_0+h}(\omega)-B_{t_0}(\omega)}{ht}\right|\leqslant C\right)=P\left(\omega:\left|\frac{B_h(\omega)}{\sqrt{h}}\right|\leqslant C\sqrt{h}\right)$$
$$=\Phi(C\sqrt{h})-\Phi(-C\sqrt{h})\to 0\quad(h\to 0),$$

其中 $\Phi(x)$ 是标准正态分布的分布函数

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t.$$

这说明：Brown 运动在任一点 t_0 处存在有限导数的概率为 0(对于我们的 Brown 运动的数学模型而言，经过概率理论的严格分析，可以进一步得到，除了概率为零的轨道以外，在每条轨道上的任意点上，其导数均不存在).

4.4　Brown 运动的首达时的分布密度

4.4.1　首达时与 Brown 运动的反射原理

对于 Brown 运动，最为重要的是**首达时** T_a，它表示 Brown 运动首次到达(或击中)a 的时间，即

$$T_a = \min\{t > 0: B_t = a\}.$$

显然,T_a 是一随机变量. 下面我们给出其分布密度.

引理 4.13 对于 $a>0$ 有 $P(B(t)\geqslant a \mid T_a\leqslant t)=\frac{1}{2}$.

直观"证明"

$$\begin{aligned} P(B(t) \geqslant a \mid T_a \leqslant t) &= P(B(t) \geqslant a \mid \text{存在 } s < t \text{ 使 } B(s) = a) \\ &\xlongequal{\text{直观地}} P(B(t) - B(s) \geqslant 0 \mid \text{存在 } s < t \text{ 使 } B(s) = a) \\ &\xlongequal{\text{独立增量性}} P(B(t) - B(s) \geqslant 0) = \frac{1}{2}. \end{aligned}$$

(这里的"推理"只是一个直观的解释,是为了使使用者能接受这个结论. 这个"推理"实际上不是合理的,因为这里的 s 不是常数,而是随机的. 因此,从严格的数学角度,这个推理是不完全正确的).

推论 4.14(Brown 运动的反射原理——对镜面的统计对称性)

$$P(B(t) \geqslant a \mid T_a \leqslant t) = P(B(t) \leqslant a \mid T_a \leqslant t),$$

P(Brown 运动在$[0,t]$到达时 a, $B_t\leqslant a$)$=P$(Brown 运动在$[0,t]$到达时 a, $B_t\geqslant a$).

证明 如图 4.3 所示. 由引理 4.13,二者都等于$\frac{1}{2}$.

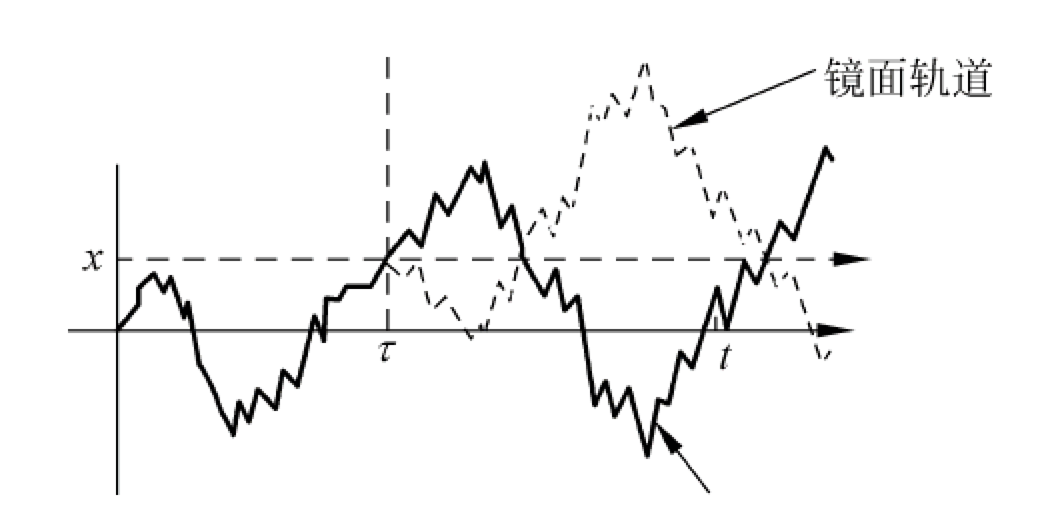


图 4.3 Brown 运动的镜面轨道

定理 4.15 对于 $a>0$,Brown 运动$\{B_t, t\geqslant 0\}$的首达 a 的首达时 T_a 的分布为

$$P(T_a \leqslant t) = 2P(B(t) \geqslant a) = 2\left[1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right].$$

证明 利用全概率公式及 $P(B(t)\geqslant a \mid T_a>t)=0$,便得到

$$\begin{aligned} P(B(t) \geqslant a) &= P(B(t) \geqslant a \mid T_a \leqslant t)P(T_a \leqslant t) + P(B(t) \geqslant a \mid T_a > t)P(T_a > t) \\ &= P(B(t) \geqslant a \mid T_a \leqslant t)P(T_a \leqslant t) = \frac{1}{2}P(T_a \leqslant t). \end{aligned}$$

推论 4.16

$$P(T_a \leqslant t) = \sqrt{\frac{2}{\pi}}\int_{\frac{a}{\sqrt{t}}}^{+\infty} \mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x.$$

因此,随机变量 T_a 在 t 的密度是

$$f_{T_a}(t)=\frac{a}{\sqrt{2\pi t^3}}e^{-\frac{a^2}{2t}}\quad(t>0).$$

推论 4.17

$$P(T_a<+\infty)=1,$$

其含义为:无论 a 多大,Brown 运动总是以概率 1 在有限时间内到达 a.

证明 $P(T_a<+\infty)=\lim\limits_{t\to+\infty}P(T_a\leqslant t)=\lim\limits_{t\to+\infty}2\left[1-\Phi\left(\frac{a}{\sqrt{t}}\right)\right]=1.$

推论 4.18

$$ET_a=+\infty,$$

其含义为:虽然 Brown 运动几乎所有的轨道首次过 a 的时间都有限,但是,过 a 点的平均时间却是无穷.

证明

$$ET_a=\int_0^{+\infty}tf_{T_a}(t)\mathrm{d}t=\int_0^{+\infty}\frac{a}{\sqrt{2\pi t}}e^{-\frac{a^2}{2t}}\mathrm{d}t=+\infty.$$

推论 4.19 对于 $a<0$ 有

$$P(T_a\leqslant t)=2\left[1-\Phi\left(\frac{-a}{\sqrt{t}}\right)\right].$$

证明 $B_0=0$,用对称性即得.

4.4.2 Brown 运动在时间区间[0,t]中达到的最大值的分布

与首达时有紧密联系的是 Brown 运动在时间区间$[0,t]$中达到的最大值$\max\limits_{0\leqslant s\leqslant t}B_s$. 显然它也是随机变量. 由 Brown 运动轨道 $B_t(\omega)$对 t 的连续性得到

$$\{\max_{0\leqslant s\leqslant t}B_s\geqslant a\}=\{T_a\leqslant t\}.$$

为了使符号更为简洁,记

$$M(t)=\max_{0\leqslant s\leqslant t}B(s).$$

即对于 $m>0$ 有

$$\{M(t)\geqslant m\}=\{T_m\leqslant t\}.$$

于是有下面的定理.

定理 4.20 当 $m>0$ 时有

$$P(M(t)\geqslant m)=P(T_m\leqslant t)=2\left(1-\Phi\left(\frac{m}{\sqrt{t}}\right)\right).$$

因此,$M(t)$的密度函数为

$$f_{M(t)}(m) = \frac{2}{\sqrt{2\pi t}} e^{-\frac{m^2}{2t}} \quad (m > 0),$$

$$EM(t) = \sqrt{\frac{2t}{\pi}}.$$

证明 函数 $2\left(1-\Phi\left(\frac{m}{\sqrt{t}}\right)\right)$对 m 的导数的负值就是 $M(t)$的密度函数在 m 处的值.

例 4.1 假定随机过程 $\xi_t=\sigma B_t$. 对于 $s<t$,有

(1)

$$\begin{aligned}P(\xi_t < 2\sigma \mid \xi_s = \sigma) &= P(\xi_t - \xi_s < \sigma \mid \xi_s = \sigma) \\ &= P(\xi_t - \xi_s < \sigma) = P\left(\frac{\xi_t - \xi_s}{\sigma\sqrt{t-s}} < \frac{\sigma}{\sigma\sqrt{t-s}}\right) = \Phi\left(\frac{1}{\sqrt{t-s}}\right).\end{aligned}$$

(2) 利用 $B(s)|_{B(t)=y} \sim N\left(\frac{s}{t}y, \frac{s}{t}(t-s)\right)$,有

$$\begin{aligned}P(\xi_s < 2\sigma \mid \xi_t = \sigma) &= P(B_s < 2 \mid B_t = 1) \\ &= P\left(N\left(\frac{s}{t}, \frac{s}{t}(t-s)\right) < 2\right) = \Phi\left(\frac{2-\frac{s}{t}}{\sqrt{\frac{s}{t}(t-s)}}\right).\end{aligned}$$

例 4.2 对于 Brown 运动,则 $P(T_1 \leqslant 1)=2[1-\Phi(1)]=0.3174$,$T_1$ 的 90%分位点 p 为方程

$$0.9 = P(T_1 \leqslant p) = 2\left[1 - \Phi\left(\frac{1}{\sqrt{p}}\right)\right]$$

的解 $p \approx 63.166$.

例 4.3 假定随机过程 $\xi_t=2B_t$,那么

(1) $P(\max\limits_{0\leqslant t\leqslant 1}\xi_t>2)=P(\max\limits_{0\leqslant t\leqslant 1}B_t>1)$

$=P(M(1)>1)=2[1-\Phi(1)]=0.3174.$

(2) $\max\limits_{0\leqslant t\leqslant 1}\xi_t$ 的中位数 m 是方程

$$\frac{1}{2} = P(\max_{0\leqslant t\leqslant 1}\xi_t \geqslant m)$$

的解,即

$$\frac{1}{2} = P\left(M(1) \geqslant \frac{m}{2}\right) = 2\left[1 - \Phi\left(\frac{m}{2}\right)\right].$$

于是有 $m=1.35$.

(3) $E(\max\limits_{0\leqslant t\leqslant 1}\xi_t)=2E(\max\limits_{0\leqslant t\leqslant 1}B(t))=2EM(1)=2\sqrt{\frac{2}{\pi}}=1.596.$

如同对称简单随机徘徊一样,选样定理对于研究 Brown 运动的首达时是一个重要的数学工具.

4.5 Brown 运动的离散近似

4.5.1 用对称随机徘徊近似

Brown 运动可以作如下的随机模拟:设$\{Y_k\}$独立同分布,且

$$P(Y_k=-1)=P(Y_k=1)=\frac{1}{2}.$$

对于固定的 t,定义如下的以 Δt 为时间单位,步长为 Δx 的对称随机徘徊

$$\eta_t^{\Delta}=\Delta x\left(Y_1+Y_2+\cdots+Y_{\left[\frac{t}{\Delta t}\right]}\right),$$

其中方括号表示取整数运算,则有

$$E\eta_t^{\Delta}=0,\quad \operatorname{var}(\eta_t^{\Delta})=(\Delta x)^2\left[\frac{t}{\Delta t}\right].$$

取充分小的数 $h>0$,并取 $\Delta t=h$,$\Delta x=\sqrt{h}$. 那么,当 $h\to 0$ 时,就有

$$\operatorname{var}(\eta_t^{\Delta})\to t.$$

由中心极限定理推出随机变量 η_t^{Δ} 的近似分布为 $N(0,t)$. 如果还用多维的中心极限定理,就可以证明随机过程$\{\eta_t^{\Delta}: t\geqslant 0\}$的所有的有限维分布,都收敛于 Brown 运动的相应的有限维分布(在数学味更为浓的随机过程课程中还证明了:在 $h\to 0$ 时,η_t^{Δ} 作为随机过程整体在某种意义下的极限是 Brown 运动). 于是随机过程$\{\eta_t^{\Delta}: t\geqslant 0\}$可以在理论上和在模拟计算中作为 Brown 运动的近似. 这是所有生成 Brown 运动的统计软件与数学软件的基本点. 如图 4.4 所示.

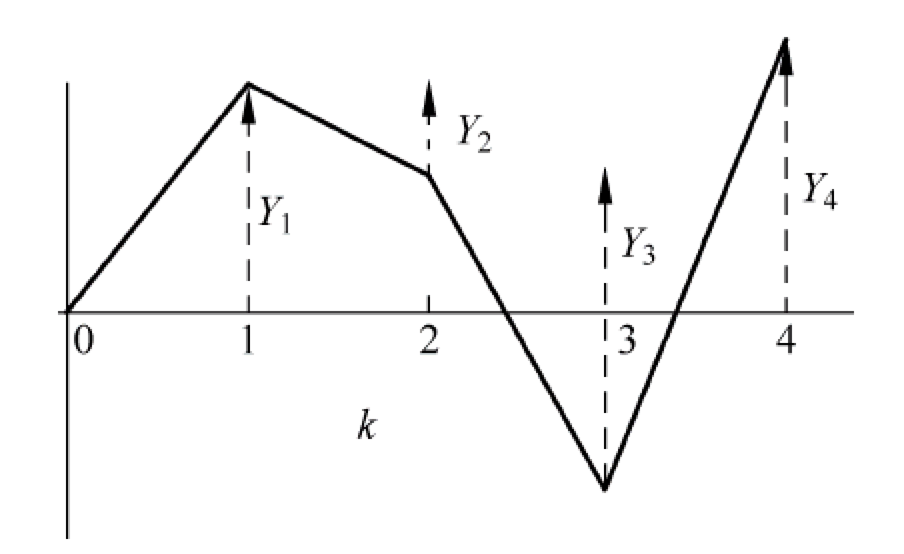


图 4.4 h 给定时,Brown 运动的近似

4.5.2 Brown 运动在击中 $b(b<0)$前,先击中 $a(a>0)$的概率

命题 4.21 对于 $a>0$,$b<0$,有

$$P(T_a < T_b) = \frac{|b|}{|b|+a}.$$

证明 由第 3 章知道，当甲有资金 A，乙有资金 $-B$ 时，甲破产的概率为 $\frac{|B|}{A+|B|}$. 也就是说，从 A 点出发的对称简单随机徘徊在击中 $A+|B|$ $(B<0)$前，先击中 0 的概率为 $\frac{|B|}{A+|B|}$. 又显见对称简单随机徘徊的概率，对于地点的平移是不变的，故而，从原点出发的对称随机徘徊，在击中 $B(B<0)$前，先击中 $A(A>0)$的概率也应为 $\frac{|B|}{A+|B|}$. 将 Brown 运动看成 $\Delta x=\sqrt{\Delta t}$ 的随机徘徊的极限，就得到

$$P(T_a < T_b) = \lim_{\Delta t\to 0}\frac{\left[\frac{-b}{\Delta x}\right]}{\left[\frac{-b}{\Delta x}\right]+\left[\frac{a}{\Delta x}\right]} = \frac{|b|}{|b|+a}.$$

4.6 Brown 运动的变种

4.6.1 漂移 Brown 运动

$\{\xi_t=\mu t+\sigma B_t:\ t\geqslant 0\}$称为分别以 μ,σ^2 为**漂移系数、扩散系数的漂移 Brown 运动**，简称为漂移 Brown 运动. 显见这时有

(1) $\xi_0=0$;

(2) $\{\xi_t:\ t\geqslant 0\}$是独立增量过程，且具有平稳增量；

(3) 对于任意 t，随机变量 $\xi_t\sim N(\mu t,\sigma^2 t)$.

可以用随机徘徊近似漂移 Brown 运动. 实际上漂移 Brown 运动可以作如下的随机模拟：

设$\{Y_k\}$独立同分布，$P(Y_k=1)=p$，$P(Y_k=-1)=1-p$. 定义 $\xi_t^{\Delta}=\sigma\Delta x(Y_1+\cdots+Y_{[\frac{t}{\Delta t}]})$，则

$$E\xi_t^{\Delta} = \Delta x\left[\frac{t}{\Delta t}\right](2p-1),\quad \operatorname{var}(\xi_t^{\Delta}) = (\sigma\Delta x)^2\left[\frac{t}{\Delta t}\right][1-(2p-1)^2].$$

对于充分小的数 $h>0$，如果取 $\Delta t=h$，$\Delta x=\sqrt{h}$，并且对于给定的常数 μ，选取 p 使

$$2p-1=\mu\sqrt{h}\quad\left(\text{即 } p=\frac{1}{2}(\mu\sqrt{h}+1)\right),$$

那么当 $h\to 0$ 时，有

$$E\xi_t^{\Delta}\to\mu t,\quad \operatorname{var}(\xi_t^{\Delta})\to\sigma^2 t.$$

由中心极限定理得到 ξ_t^{Δ} 的近似分布为 $N(\mu t,\sigma^2 t)$. 同样，如果用多维的中心极限定理，就可以证明随机过程$\{\xi_t^{\Delta}:\ t\geqslant 0\}$的所有的有限维分布，都收敛于漂移 Brown 运动的相应的

有限维分布.进一步可以得到：在 $h\to0$ 时，$\{\xi_t^\Delta: t\geqslant0\}$作为随机过程整体在某种意义下的极限是漂移 Brown 运动.于是随机过程$\{\xi_t^\Delta: t\geqslant0\}$的任何泛函都可以近似漂移 Brown 运动的相应泛函.

4.6.2 几何 Brown 运动

如果$\{\xi_t: t\geqslant0\}$是漂移 Brown 运动，则 $\zeta_t=e^{\xi_t}$ 称为**几何 Brown 运动**.对于几何 Brown 运动，有

(1) $\zeta_0=1$；

(2) $\zeta_t=e^{\xi_t}$ 服从均值为 μt，方差为 $\sigma^2 t$ 的对数正态分布.又若 $s+t\leqslant u$，则$\frac{\zeta_{s+t}}{\zeta_s}$与$\frac{\zeta_{u+t}}{\zeta_u}$独立，且都与 $\zeta_t=e^{\xi_t}$ 同分布.

命题 4.22

$$E\zeta_t^k = Ee^{\xi_t k} = e^{(\mu t)k+\frac{(\sigma^2 t)k^2}{2}},$$

从而推出

$$E\zeta_t = e^{\mu t+\frac{\sigma^2 t}{2}},$$
$$\mathrm{var}(\zeta_t) = e^{2\mu t+\sigma^2 t}(e^{\sigma^2 t}-1).$$

证明 $Ee^{\xi_t k}$ 是正态随机变量 ξ_t 的矩母函数在 k 处的值，即

$$e^{(E\xi_t)k+\frac{\mathrm{var}(\xi_t)k^2}{2}} = e^{(\mu t)k+\frac{(\sigma^2 t)k^2}{2}}.$$

例 4.4 在著名的 Black-Scholes 模型中，人们用随机变量 $P_t=P_0\zeta_t$ 描述一种证券在正常时刻 t 的价值，并以 $e^{-\delta t}P_t$ 表示其在开始时刻 0 的折现价.假定 $\mu+\frac{\sigma^2}{2}=\delta$，那么折现价$\{e^{-\delta t}P_t: t\geqslant0\}$是$\{B_t\}$鞅.

证明
$$\begin{aligned}E[e^{-\delta(t-s)}P_t \mid \boldsymbol{B}_s] &= P_s e^{-\delta(t-s)}E\left(\frac{P_t}{P_s}\middle|\boldsymbol{B}_s\right)\\ &= P_s e^{-\delta(t-s)}(e^{\mu(t-s)}e^{\sigma(B_t-B_s)} \mid \boldsymbol{B}_s) = P_s e^{(\mu-\delta)(t-s)}Ee^{\sigma B_{t-s}}\\ &= P_s e^{(\mu-\delta)(t-s)}e^{\frac{\sigma^2(t-s)}{2}} = P_s e^{\left(\mu+\frac{\sigma^2}{2}-\delta\right)(t-s)} = P_s.\end{aligned}$$

所以

$$E[e^{-\delta t}P_t \mid \boldsymbol{B}_s] = e^{-\delta s}P_s.$$

例 4.5 假定 100 份证券，价格过程用几何 Brown 运动建模，每份证券的初始价格为 $P_0=100$，$\delta=\mu=\sigma^2=0.02$.那么，在时刻 $t=1$ 有

(1) $EP_1=100e^{\mu+\frac{\sigma^2}{2}}=100e^{0.03}=103.05$，

$$\mathrm{var}(P_1) = 10000\mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2}-1) = 10000\mathrm{e}^{0.06}(\mathrm{e}^{0.02}-1) = 214.51.$$

(2) $P(P_1>100\mathrm{e}^{0.03}) = P(\xi_1>0.03)$

$$= P\left(N(0,1)>\frac{0.03-0.02}{\sqrt{0.02}}\right) = 1-\Phi(0.0707) = 0.4782.$$

(3) $P(\max\limits_{0\leqslant t\leqslant 5}\mathrm{e}^{-0.02t}P_t>100\mathrm{e}^{0.04}) = P(\max\limits_{0\leqslant t\leqslant 1}\mathrm{e}^{-0.02t}\mathrm{e}^{0.02t+\sqrt{0.02}B_t}>\mathrm{e}^{0.04})$

$$= P\left(M_5>\frac{0.04}{\sqrt{0.02}}\right) = P(M_5>0.2828)$$

$$= 2\left[1-\Phi\left(\frac{0.2828}{\sqrt{5}}\right)\right] = 2[1-0.5503] = 0.899.$$

(4) $P(\mathrm{e}^{-\delta t}P_t$ 在到达 $100\mathrm{e}^{-0.01}$ 前先到达 $100\mathrm{e}^{0.02})$

$$= P\left(B_t\text{ 在到达}\frac{-0.01}{\sqrt{0.02}}\text{前先到达}\frac{0.02}{\sqrt{0.02}}\right) = \frac{0.01}{0.01+0.02} = \frac{1}{3}.$$

例 4.6(Black-Scholes 证券模型的欧式期权的定价)　设某人拥有一份在将来的时刻 T(**执行日**, executive time)以**敲定价格**(striking price)K, 购买 100 股某种证券的**欧式看涨期权**(European call option). 假设该 100 股目前的价格为 y, 且它的随机价格过程 η_t 按照几何 Brown 运动变化, 即 $\eta_t = y\mathrm{e}^{\mu t+\sigma B_t}$. 若到了执行日 T, 该 100 股的价格高于敲定价格 K, 那么, 他就实施该期权. 否则, 他可以放弃实施该期权. 因此, 在执行日 T 他拥有的期权的平均价值为

$$E\max\{\eta_T-K,0\} = \int_0^{+\infty}P(\eta_T-K>a)\mathrm{d}a = \int_0^{+\infty}P(y\mathrm{e}^{\mu T+\sigma B_T}-K>a)\mathrm{d}a$$

$$= \int_0^{+\infty}P\left(\mu T+\sigma B_T>\ln\frac{a+K}{y}\right)\mathrm{d}a.$$

注意 $\mu T+\sigma B_T\sim N(\mu T,\sigma^2 T)$, 故上式等于

$$\frac{1}{\sigma\sqrt{2\pi T}}\int_0^{+\infty}\left[\int_{\ln\frac{a+K}{y}}^{+\infty}\mathrm{e}^{-\frac{(x-\mu T)^2}{2\sigma^2 T}}\mathrm{d}x\right]\mathrm{d}a.$$

假定银行的连续利率为 r. 那么, 此期权在开始时间的价格应是在执行日 T 的折现, 即 $\mathrm{e}^{-rT}E[\max\{\eta_T-K,0\}]$.

例 4.7　假定 $\{\xi_t: t\geqslant 0\}$ 是漂移系数与扩散系数分别为 μ,σ^2 的漂移 Brown 运动. 求在 $\xi_1=\mu+0.1\sigma$ 条件下, $\xi_{\frac{1}{2}}$ 的条件分布?

解　因为条件 $\xi_1=\mu+0.1\sigma$ 等价于 $B_1=0.1$, $\xi_{\frac{1}{2}}=\frac{1}{2}\mu+\sigma B_{\frac{1}{2}}$. 而

$$B_{\frac{1}{2}}\Big|_{B_1=0.1}\sim N\left(\frac{1}{2}\times 0.1,\frac{1}{2}\left(1-\frac{1}{2}\right)\right) = N\left(0.05,\frac{1}{4}\right),$$

于是

$$\xi_{\frac{1}{2}}\Big|_{\xi_1=\mu+0.1\sigma}=\left(\frac{1}{2}\mu+\sigma B_{\frac{1}{2}}\right)\Big|_{B_1=0.1}\sim N\left(\frac{\mu}{2}+0.05\sigma,\frac{\sigma^2}{4}\right).$$

例 4.8 假定某个证券的价格是时间以月为单位的标准 Brown 运动，其初始价格为 18. 求在此后的 4 个月中，其价格曾到达 21 的概率.

解 价格的波动幅度为 21－18＝3. 令 T_3 为标准 Brown 运动初达 3 的时刻. 那么，所要的概率就是

$$P(T_3\leqslant 4)=2\left[1-\Phi\left(\frac{3}{\sqrt{4}}\right)\right]=2[1-0.9332]=0.1336.$$

例 4.9 假定目前 A 国货币与 B 国货币等值，而在时刻 t 两国货币的差为 $C(t)$，即在时刻 t，1 元 A 币兑换 $1+C(t)$ 元 B 币. 而 $C(t)$ 是标准 Brown 运动的 0.01 倍. A 国的某投资人在 B 国的无风险环境下投资 1 元，在 5 年后增值为 B 币 1.5 元. 已知 1 年后 A 币 1 元等值于 B 币 1.05 元. 求在此条件下，在到期时此资金至少值 1.5 元 A 币的条件概率.

解 条件是 $C(1)=1.05-1=0.05$. 在投资 5 年到期时，资金的 B 币值为 1.5，而其 A 币值就是$\frac{1.5}{1+C(5)}$. 于是所求的条件概率为

$$\begin{aligned}P\left(\frac{1.5}{1+C(5)}\geqslant 1.5 \mid C(1)=0.05\right)&=P(C(5)\leqslant 0\mid C(1)=0.05)\\&=P(C(5)-C(1)\leqslant -0.05\mid C(1)=0.05)\\&\xlongequal{\text{独立增量性}}P(C(5)-C(1)\leqslant -0.05)\\&=P(C(4)\leqslant -0.05)\\&=P\left(\frac{C(4)}{\sqrt{0.04}}\leqslant -0.25\right)=1-\Phi(0.25)=0.4.\end{aligned}$$

4.6.3 0 点反射 Brown 运动

$B_t^{(r)}\xlongequal{\text{def}}|B_t|$ 称为 **0 点反射 Brown 运动**. 利用 Gamma 函数的性质，可以得到

$$E|B_t|=\sqrt{\frac{2t}{\pi}},\quad \operatorname{var}|B_t|=\left(1-\frac{2}{\pi}\right)t.$$

4.6.4 积分 Brown 运动

$\xi_t\xlongequal{\text{def}}\int_0^t B_s\,\mathrm{d}s$ 称为**积分 Brown 运动**，它的近似积分和是 Gauss 过程，因此积分 Brown 运动作为 Gauss 过程的极限，也是 Gauss 过程，满足

$$E\left(\int_0^t B_s\,\mathrm{d}s\right)=0,$$

$$\mathrm{cov}\left[\int_0^s B_u \mathrm{d}u, \int_0^t B_v \mathrm{d}v\right] = \int_0^s \int_0^t (u \wedge v) \mathrm{d}u \mathrm{d}v.$$

积分 Brown 运动的轨道不仅连续,而且按轨道具有连续导数,其按轨道的导数随机过程就是 Brown 运动. 积分 Brown 运动不再具有 Markov 性. 但是不难证明,二元随机过程 $\left(\int_0^t B_s \mathrm{d}s, B_t\right)$ 却是二维的 Markov 过程.

习题 4

1. 对于一维 Brown 运动 B_t,证明当 $t \to +\infty$ 时,$\dfrac{B_t}{t}$ 依概率收敛到 0.

2. 求 $\xi_t = \mathrm{e}^{-t} B_{\mathrm{e}^{2t}}$ 的协方差函数.

3. 求一维积分 Brown 运动 $\xi_t = \int_0^t B_s \mathrm{d}s$ 的方差与协方差函数.

4. 设 B_t 为 Brown 运动,分别求(1)$\xi_t = \sigma B_t + \mu t$,(2)$\eta_t = (1-t) B_{\frac{t}{1-t}} (t < 1)$,(3)$\zeta_t = \mathrm{e}^{-\beta t} B_{\mathrm{e}^{2\beta t}}$ 的协方差函数.

5. 设 $s < u < t$,B_t 为 Brown 运动. 求 $E(B_s \mid B_t)$,$E(B_t \mid B_s)$,$E(B_t \mid B_s, B_u)$,$E(B_s \mid B_u, B_t)$,$E(B_u \mid B_s, B_t)$.

6. 证明 Brown 运动有空间平移不变性,即

$$P(B_{t+s} \in A \mid B_s = x) = P(B_{t+s} \in A + z \mid B_s = x + z),$$

其中 $A+z=\{y+z:\ y\in A\}$.

7. 设 $a<0$,对于 Brown 运动 $B_t(B_0=0)$,给出在 a 点的反射原理及吸附的 Brown 运动的概率规律.

8. 对于 $0<t<1$,证明在条件 $B_0=B_1=0$ 条件下,B_t 的条件分布密度为 $N(0,t(1-t))$. 再对 $B_0=a, B_1=b$ 的情形推广这个结果.

9. 对于 $0\leqslant t_1<t_2<\cdots<t_{n+1}$. 求条件分布 $P\{B(t_{n+1})\leqslant x \mid B(t_1)=x_1,\cdots,B(t_n)=x_n\}$ 和 $P\{B(t_1)\leqslant x \mid B(t_2)=x_2,\cdots,B(t_{n+1})=x_{n+1}\}$ 的密度,并由此再求相应的条件期望和条件方差.

10. 求 $\xi_t=B_t-tB_1$ 及 $\eta_t=(t+1)\xi_{\frac{t}{t+1}}$ 的分布. 再对 $0<s<t<1$ 求 ξ_s 与 ξ_t 的协方差函数与相关系数.

11. 分别求 $|B_t|$,B_t^*,$B_{*t}=\min\limits_{s\leqslant t} B_s$ 及 $D_t=B_t^*-B_t$ 的密度. 求证

$$P\{B_t^* > x \mid D_t = 0\} = \exp(-x^2/2t).$$

12. 记 $S_n = \sum\limits_{k=1}^{2^n} (B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}})^2$. 求 ES_2,$E(S_3 \mid S_2)$,$E(S_2 \mid S_3)$,ES_n,并证明

$$E(S_{n+1} \mid S_n) = \frac{1}{2}(S_n + 1), \quad E(S_n \mid S_{n+1}) = S_{n+1}.$$

13. 对于漂移 Brown 运动,求$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$在点$(x_1,x_2,\cdots,x_n)$的联合密度.

14. 对于漂移 Brown 运动,$t_1<t_2$,求条件分布 $\xi_{t_2}|_{B_{t_1}}$ 与 $\xi_{t_1}|_{B_{t_2}}$.

15. 对于漂移 Brown 运动,$s<t$,求 $P(\xi_t<2\sigma|\xi_s=\sigma)$与 $P(\xi_s<2\sigma|\xi_t=\sigma)$.

16. 怎样的 Gauss 过程是漂移 Brown 运动?

17. 证明 Brown 桥 $X_t=B_t-tB_1(0\leqslant t\leqslant 1)$是 Markov 过程.

18. 对于积分 Brown 运动 $\xi_t=\int_0^t B_s\mathrm{d}s$ 证明(B_t,ξ_t)是二维 Markov 过程.

第 5 章　随机微积分，对 Brown 运动的 Ito 积分与 Ito 公式

通常所说的随机微积分，就是对 Brown 运动的积分与由它发展出来的 Ito 过程及 Ito 公式. 因为其定义是由日本数学家 Ito 给出的，所以称为 Ito 积分. Ito 积分是随机分析的基本工具.

随机微积分的内容，涉及较深的数学，要真正表达清楚，需要测度论等较为严格的数学工具. 我们不希望过多地使用严格数学. 因此，我们叙述定义与定理时所要求的条件不求精确，而更多地关注于用典型例子开路给出其想法，以求随机微积分的思想精髓.

随机微积分的对象是一种特殊的随机函数，即 Ito(随机)过程. 大家知道，在普通微积分中，复合函数的微积分是函数研究的精髓. 而 Ito 过程的复合函数，也是随机微积分的主旋律. 普通函数的微分与积分是一对互逆的运算. 对于较为简单的函数，人们常以函数的导数作为起点，进一步得到微分与积分的运算与理论. 然而，事实上对于复杂的函数，积分却比导数更容易理解，处理与推广，而且更容易于近似计算. 由此人们认识到，以积分作为微积分的核心是更为合理的. 而随机微积分正是以对 Brown 运动的积分，即 Ito 积分，作为核心发展的.

为了了解对 Brown 运动的积分与普通积分的本质不同，我们先要叙述普通定积分的一种简单推广，就是 Stieltjes 积分. 然而，我们并不能对每一条 Brown 轨道 $B_t = B_t(\omega)$ ($\omega \in \Omega$ 固定)定义 Stieltjes 积分，而要用 Ito 的办法，将 Brown 运动作为随机函数，对它整体地定义积分.

5.1　实值函数的 Stieltjes 积分

5.1.1　对单调函数的 Stieltjes 积分

命题 5.1　设 $F(t)$ 为区间 $[a,b]$ 上的单调递增函数，$g(t)$ 为连续函数. 对于 $[a,b]$ 的一个划分

$$a = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n}^{(n)} = b,$$

$g(t)$ 对 $F(t)$ 的 **Riemann-Stieltjes** 和定义为

$$\sum_{i=0}^{N_n-1} g(\xi_i)[F(t_{i+1}^{(n)})-F(t_i^{(n)})], \quad t_i \leqslant \xi_i \leqslant t_{i+1}.$$

那么,当划分的最大半径

$$\lambda_n = \max_i\{t_{i+1}^{(n)} - t_i^{(n)}\}$$

趋于 0 时,可以证明 Stieltjes 和具有不依赖划分和 ξ_i 们的选取的极限,将这个极限记为 $\int_a^b g(t)\mathrm{d}F(t)$,称为 $g(t)$对 $F(t)$的 **Stieltjes 积分**.

假定 $F(t)$是随机变量 ξ 的分布函数. 如果

$$\int |t| \mathrm{d}F(t) \xlongequal{\text{def}} \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b |t| \mathrm{d}F(t)$$

存在,则

$$\int t\mathrm{d}F(t) \xlongequal{\text{def}} \lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b t\mathrm{d}F(t)$$

也存在,这时有

$$E\xi = \int t\mathrm{d}F(t).$$

如果 $F(t)$有导数 $f(t)=\dfrac{\mathrm{d}F}{\mathrm{d}t}$,且$|f(t)|$可积,那么,Stieltjes 积分就可以转化为普通的定积分

$$\int_a^b g(t)\mathrm{d}F(t) = \int_a^b g(t)f(t)\mathrm{d}t.$$

如果 $g(t)$连续可微,那么,Stieltjes 积分的计算也可以转化为定积分计算:

$$\int_a^b g(t)\mathrm{d}F(t) = [g(b)F(b)-g(a)F(a)] - \int_a^b F(t)g'(t)\mathrm{d}t.$$

(将两个积分的积分和相加后简化,取极限便得).

5.1.2 对有界变差函数的 Stieltjes 积分

为了弄清楚 Stieltjes 积分中 $F(t)$的应用范围,我们介绍有界变差函数的概念如下.

定义 5.1 对于区间$[a,b]$如下的一个划分 $P_{t_0^{(n)},\cdots,t_{N_n}^{(n)}}$:

$$a = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n}^{(n)} = b,$$

我们将

$$V(P_{t_0^{(n)},\cdots,t_{N_n}^{(n)}}) \xlongequal{\text{def}} |f(t_1^{(n)}) - f(t_0^{(n)})| + \cdots + |f(t_{N_n}^{(n)}) - f(t_{N_{n}-1}^{(n)})|$$

称为函数 $f(t)$在划分上的变差,并将 $f(t)$在所有的划分上的变差的上确界:

$$V_{[a,b]}(f) \xlongequal{\text{def}} \sup_{n,P_{t_0^{(n)},\cdots,t_{N_n}^{(n)}}} V(P_{t_0^{(n)},\cdots,t_{N_n}^{(n)}})$$

称为函数 $f(t)$在区间$[a,b]$上的**全变差**.

在区间$[a,b]$上全变差为有限值的函数,称为在区间$[a,b]$上的**有界变差函数**(bounded variation function),或有界变化函数.

在不会混淆的情形下,$V_{[a,b]}(f)$常将简单地记为$V(f)$.从定义可以直接验证,如果$a<c<b$,则有

$$V_{[a,b]}(f) = V_{[a,c]}(f) + V_{[c,b]}(f).$$

有关有界变差函数的基本事实如下.

(1) 关于有界变差函数结构的基本定理是:函数是有界变差函数的充分必要条件是,它是两个单调增函数的差.这里的单调增函数在不计一个常数加项外是唯一的.

单调函数是(从而有界变差函数也是)几乎处处可导的,即它的不可微的点的全体,记为N_0,是在下述意义下的"零测度"集合:

对于任给$\varepsilon>0$,存在区间列$\{I_i,i\geqslant 1\}$,使$\bigcup\limits_{i=1}^{+\infty} I_i \supset N_0$,且$I_i$的长度小于$\frac{\varepsilon}{2^{i+1}}$.

(2) 连续函数一般并不是有界变差函数.反之,有界变差函数一般也并不一定连续.

(3) 最简单的有界变差函数是某个区间上的示性函数.

(4) 两个有界变差函数的线性组合是有界变差函数.

因为$F(t)$是有界变差函数总能表示为两个单调递增函数的差

$$F(t) = F_1(t) - F_2(t),$$

所以可以定义 Stieltjes 积分

$$\int_a^b g(t)\mathrm{d}F(t) = \int_a^b g(t)\mathrm{d}F_1(t) - \int_a^b g(t)\mathrm{d}F_2(t).$$

注 1　当$F(t)=t$时,Stieltjes 积分就是普通微积分中的定积分.

注 2　如果不假定函数$g(t)$的连续性,也不假定$F(t)$的单调性,当划分的最大半径

$$\lambda_n = \max_i\{t_{i+1}^{(n)} - t_i^{(n)}\}$$

趋于0时,只要 Riemann-Stieltjes 和具有不依赖划分和ξ_i们的选取的极限,则称$g(t)$对$F(t)$Stieltjes 可积,并称此极限为$g(t)$对$F(t)$的 Stieltjes 积分.

Stieltjes 积分的存在性理论,是与定积分的存在性理论完全类似的.粗略地讲,为了使$g(t)$对$F(t)$Stieltjes 可积,$F(t)$在每个局部不能有无界的变差,所以,在一般情形,$F(t)$的有界变差性几乎是不可缺的.但是,如果$g(t)$连续可微,那么,对于任意连续函数$F(t)$,我们可以通过定积分用如下的公式,来定义$g(t)$对$F(t)$的 Stieltjes 积分:

$$\int_a^b g(t)\mathrm{d}F(t) \stackrel{\text{def}}{=\!=} [g(b)F(b) - g(a)F(a)] - \int_a^b F(t)g'(t)\mathrm{d}t.$$

事实上,这样的补充定义的意义并不太大,因为定义 Stieltjes 积分的目的,本来就是要处理一些普通的定积分不能处理的问题,既然定积分已经能表达了,再写成 Stieltjes 积分就只是形式的不同,有时只是看似简单一些而已.

5.2　对 Brown 运动的随机积分

引理 5.2　对 Brown 运动 $\{B_t: t\geqslant 0\}$ 有

$$P(\omega: B_t(\omega)\text{ 是 } t \text{ 的处处连续,且处处不可微函数}) = 1.$$

证明　只需考虑 $t\in[0,1]$ 情形. 假设对于固定的 ω,函数 $B_t(\omega)$ 在 $t=s(\in(0,1))$ 可微,那么存在 m_0 及 l,只要 $m\geqslant m_0$ 以及 $|t-s|<\dfrac{1}{m}$,就有 $|B_t(\omega)-B_t(\omega)|<l|t-s|$.

当 $n>4m$ 时,令 $i=[ns]+1$,则 $\dfrac{i-1}{n}=\dfrac{[ns]}{n}<s\leqslant\dfrac{[ns]+1}{n}=\dfrac{i}{n}$. 于是对 $j=i+1,i+2,i+3$ 有 $\left|\dfrac{j}{n}-s\right|,\left|s-\dfrac{j-1}{n}\right|<\dfrac{4}{n}<\dfrac{1}{m}$,故而

$$\begin{aligned}|B_{\frac{j}{n}}-B_{\frac{j-1}{n}}|&\leqslant|B_{\frac{j}{n}}-B_s|+|B_s-B_{\frac{j-1}{n}}|\\&\leqslant l\left|\frac{j}{n}-s\right|+l\left|s-\frac{j-1}{n}\right|\leqslant l\left(\frac{4}{n}+\frac{4}{n}\right)=8\,\frac{l}{n}.\end{aligned}$$

记

$$A_{n,l}^i=\left\{\omega: |B_{\frac{j}{n}}(\omega)-B_{\frac{j-1}{n}}(\omega)|\leqslant 7\,\frac{l}{n}, j=i+1,i+2,i+3\right\}.$$

则显然 $A_{n,l}^i$ 对于 l 递增. 并且由上面的取法得到

$$\{\omega: \exists s\in(0,1), B_t(\omega)\text{ 在点 } s \text{ 可微}\}\subset\bigcup_l\bigcap_{m\geqslant m_0}\bigcap_{n>4m}\bigcup_{i\leqslant n-3}A_{n,l}^i.$$

然而,由于

$$\begin{aligned}P\Big(\bigcap_{n>4m}\bigcup_{i\leqslant n-3}A_{n,l}^i\Big)&\leqslant P\Big(\bigcup_{i\leqslant 5m}A_{5m,l}^i\Big)\leqslant 5m\max_{i\leqslant 5m}P(A_{5m,l}^i).\\&\leqslant 5mP\left(|B_{\frac{j}{5m}}(\omega)-B_{\frac{j-1}{5m}}(\omega)|\leqslant 8\,\frac{l}{5m}\right)^3=5m\left(2\int_0^{8\frac{l}{5m}}\frac{1}{\sqrt{2\pi\,\frac{1}{5m}}}\mathrm{e}^{-\frac{u^2}{2\frac{1}{5m}}}\mathrm{d}u\right)^3\\&\leqslant 5m\left(\frac{1}{\sqrt{2\pi\,\frac{1}{5m}}}16\,\frac{l}{5m}\right)^3=C\,\frac{l^3}{\sqrt{m}}\xrightarrow{m\to+\infty}0.\end{aligned}$$

于是

$$P\Big(\bigcap_{m\geqslant m_0}\bigcap_{n>4m}\bigcup_{i\leqslant n-3}A_{n,l}^i\Big)=0.$$

从而 $P\{\omega: \exists s\in(0,1), B_t(\omega)\text{在 } s \text{ 可微}\}\leqslant P\Big(\bigcup_l\bigcap_{m\geqslant m_0}\bigcap_{n>4m}\bigcup_{i\leqslant n-3}A_{n,l}^i\Big)=0.$

读者完全可以忽略引理 5.2 的证明.

推论 5.3　对 Brown 运动$\{B_t: t\geqslant 0\}$有

$$P(\omega: B_t(\omega) \text{ 在某个有限时间区间上有界变差}) = 0.$$

证明　因为有界变差函数是两个单调函数的差，所以几乎处处可导.

这就是说，Brown 运动的样本轨道函数在任意区间上都不是有界变差的. 这一点决定了对 Brown 运动的积分的特殊性. 为了剖析得更清晰一些，下面我们从最简单的特殊情形开始定义对 Brown 运动的积分，然后逐渐地扩展它的含义.

5.2.1　实值函数对 Brown 运动的积分

如果 $f(t)$是定义在$[0,T]$上的实函数，如何定义它对 Brown 运动的积分呢？

因为 Brown 运动的样本轨道函数在任意区间上都不是有界变差的，所以即使对于一般的连续函数 $f(t)$，仍然不能对每个固定的基本事件 ω，将 $f(t)$对 Brown 运动的积分定义为对 Brown 运动的样本函数 $B_t(\omega)$的（分别按轨道的）Stieltjes 积分 $\int_0^T f(t)\mathrm{d}B_t(\omega)\left(\overset{\text{即}}{=}\int_0^T f(t)\mathrm{d}B(t,\omega)\right)$.

然而，对于连续可微而且满足 $f(0)=f(T)=0$ 的函数 $f(t)$，我们可以用下面的式子定义 $f(t)$对 Brown 运动的积分：

$$\int_0^T f(t)\mathrm{d}B_t(\omega) = -\int_0^T f'(t)B_t(\omega)\mathrm{d}t,$$

有人将这种特殊的积分称为 **Paley-Wiener-Zygmund 积分**. 它确实可以认为是“按轨道”的积分. 这样的积分是一个随机变量，除了线性性质外，它满足

(I.1) $E\left[\int_0^T f(t)\mathrm{d}B_t(\omega)\right] = 0$；

(I.2) $E\left[\int_0^T f(t)\mathrm{d}B_t(\omega)\right]^2 = \int_0^T f(t)^2\mathrm{d}t$,

$$\operatorname{cov}\left(\int_0^T f(t)\mathrm{d}B_t(\omega),\int_0^T g(t)\mathrm{d}B_t(\omega)\right) = \int_0^T f(t)g(t)\mathrm{d}t.\quad (g(t)\text{ 连续可微}).$$

证明　由于 $f'(t)$是连续函数，这里用了取期望运算与积分运算的可交换性.

(I.1) $E\left[\int_0^T f(t)\mathrm{d}B_t(\omega)\right] = -E\left[\int_0^T f'(t)B_t(\omega)\mathrm{d}t\right] = -\int_0^T f'(t)E[B_t(\omega)]\mathrm{d}t = 0.$

(I.2) $E\left[\int_0^T f'(t)\mathrm{d}B_t(\omega)\mathrm{d}t\int_0^T f'(s)\mathrm{d}B_s(\omega)\mathrm{d}s\right]$

$$= \int_0^T \mathrm{d}t\int_0^T f'(t)f'(s)E[B_t(\omega)B_s(\omega)]\mathrm{d}s = \int_0^T\int_0^T f'(t)f'(s)t\wedge s\mathrm{d}s\mathrm{d}t$$

$$= \int_0^T f'(t)\left[\int_0^t f'(s)s\mathrm{d}s + t\int_t^T f'(s)\mathrm{d}s\right]\mathrm{d}t(\text{利用 } f(0) = f(T) = 0)$$

$$= \int_0^T f'(t)\left[\int_0^t f'(s)s\mathrm{d}s - tf(t)\right]\mathrm{d}t = -\int_0^T f'(t)\int_0^t f(s)\mathrm{d}s\mathrm{d}t = \int_0^T f(t)^2\mathrm{d}t.$$

注意,如果 f_n, f 都连续可微,满足 $f_n(0)=f_n(T)=f(0)=f(T)=0$,且 $\int_0^T |f_n(t)-f(t)|^2 \mathrm{d}t \to 0$,那么,由(I.2)有

$$E\left[\int_0^T f_n(t)\mathrm{d}B_t(\omega)-\int_0^T f(t)\mathrm{d}B_t(\omega)\right]^2=\int_0^T |f_n(t)-f(t)|^2 \mathrm{d}t \to 0,$$

即 $\int_0^T f_n(t)\mathrm{d}B_t(\omega)$ 均方收敛到 $\int_0^T f(t)\mathrm{d}B_t(\omega)$,因此 $E\left[\int_0^T f_n(t)\mathrm{d}B_t(\omega)\right]^2$ 收敛到 $E\left[\int_0^T f(t)\mathrm{d}B_t(\omega)\right]^2$. (I.2)中最后的等式得自

$$\begin{aligned}&\operatorname{cov}\left(\int_0^T f(t)\mathrm{d}B_t(\omega),\int_0^T g(t)\mathrm{d}B_t(\omega)\right)\\&=\frac{1}{2}\left(E\left[\int_0^T (f(t)+g(t))^2\mathrm{d}B_t(\omega)\right]-E\left[\int_0^T f(t)^2\mathrm{d}B_t(\omega)\right]-E\left[\int_0^T g(t)^2\mathrm{d}B_t(\omega)\right]\right).\end{aligned}$$

虽然对于没有导数的函数,如上的定义显然不可行.可是以上的收敛方式仍然相当本质地揭示了随机积分间应有的极限联系.

现在我们要对连续函数 $f(t)$,定义它对 Brown 运动的积分.

对于连续函数 $f(t)$,仍旧考虑区间$[0,T]$的划分:

$$0=t_0^{(n)}<\cdots<t_{N_n}^{(n)}=T,\quad \lambda_n \overset{\text{def}}{=} \max_{0\leqslant k\leqslant N_n-1}\{t_{k+1}^{(n)}-t_k^{(n)}\}\to 0,$$

并且作 $f(t)$的梯形近似函数

$$f_n(t)\overset{\text{def}}{=}\sum_{k=0}^{N_n-1} f(t_k^{(n)}) I_{(t_k^{(n)},t_{k+1}^{(n)}]}(t),$$

(请注意,$t_k^{(n)}$ 是划分的区间的左端点)并定义如下的积分和

$$I_n(\omega)\overset{\text{def}}{=}\int_0^T f_n(t)\mathrm{d}B_t(\omega)\overset{\text{def}}{=}\sum_{k=0}^{N_n-1} f(t_k^{(n)})(B_{t_{k+1}^{(n)}}(\omega)-B_{t_k^{(n)}}(\omega)).$$

这时虽然有 $f_n(t)\to f(t)$,可是对于固定的 ω,极限 $\lim\limits_{n\to+\infty} I_n(\omega)$一般并不存在,故而我们无法直接将 $\int_0^T f(t)\mathrm{d}B_t(\omega)$ 定义为 $\lim\limits_{n\to+\infty} I_n(\omega)$.

一般性的数学思维是在不能做到存在极限时,减弱极限的含义,使得在较弱的意义下极限存在.就是说,如果我们退一步,不再要求 $\lim\limits_{n\to+\infty} I_n(\omega)$对每一个 ω 都存在,而将 $I_n \overset{\text{def}}{=} I_n(\omega)(\omega\in\Omega)$看成随机变量列,那么,它在按均方收敛的意义下(因而也依概率收敛的意义下),就可以有一个极限随机变量.为了看清楚这一点,我们只需证明,I_n 在 $L^2(\Omega)$中确实存在(从而唯一的)一个极限($L^2(\Omega)$中收敛是平均收敛,与普通收敛是不同的,但是它蕴涵了依概率收敛,也就是至少强于依概率收敛),从而可以将这个极限随机变量定义为 $f(t)$对 Brown 运动的积分,并记为 $\int_0^T f(t)\mathrm{d}B_t$. 对此,我们在下面的课文中,将给它一个较细的阐述,并显示数学工具所发挥的作用.

定义 5.2 如果$(0,T]$的子区间$(s_i,t_i]$($i=1,2,\cdots,m$)两两不相交,则称

$$g(t) \xlongequal{\text{def}} \sum_{i=0}^{m} c_i I_{(s_i,t_i]}(t)$$

为**梯形函数**. 对于任意固定的基本事件 ω，定义梯形函数对 Brown 运动的积分为

$$\int_0^T g(t)\mathrm{d}B_t(\omega) \xlongequal{\text{def}} \sum_{i=0}^{m} c_i (B_{t_i}(\omega) - B_{s_i}(\omega)).$$

显见一个梯形函数的不同表示在不计零概率事件差异下并不会影响作为随机变量的积分的"值".

$\int_0^T g(t)\mathrm{d}B_t(\omega)$ 通常简单地记为 $\int_0^T g(t)\mathrm{d}B_t$.

引理 5.4　梯形函数 $g(t)$对 Brown 运动的积分满足(这些性质相同于(I.1)与(I.2))

$$E\left[\int_0^T g(t)\mathrm{d}B_t(\omega)\right] = 0,$$

$$E\left[\int_0^T g(t)\mathrm{d}B_t(\omega)\right]^2 = \sum_{i=0}^{m} c_i^2(t_i - s_i) = \int_0^T |g(t)|^2 \mathrm{d}t.$$

后者是梯形函数$|g(t)|^2$ 的定积分.

证明　利用 Brown 运动的独立增量性质

$$\begin{aligned} E\left[\int_0^T g(t)\mathrm{d}B_t(\omega)\right]^2 &= E\left[\sum_{i=0}^{m} c_i (B_{t_i}(\omega) - B_{s_i}(\omega))\right]^2 \\ &= \sum_{i=0}^{m} c_i^2 E\left[B_{t_i}(\omega) - B_{s_i}(\omega)\right]^2 \\ &= \sum_{i=0}^{m} c_i^2(t_i - s_i) = \int_0^T |g(t)|^2 \mathrm{d}t. \end{aligned}$$

引理 5.5　对于如上定义的连续函数 $f(t)$的梯形近似函数列 $f_n(t)$，有

$$\int_0^T |f_n(t) - f(t)|^2 \mathrm{d}t \to 0.$$

证明　在有限区间$[0,T]$上，任给 $\varepsilon>0$，只要 n 充分大，就有$|f_n(t)-f(t)|<\varepsilon$. 于是

$$\int_0^T |f_n(t) - f(t)|^2 \mathrm{d}t < \varepsilon^2 T.$$

后者能充分小.

推论 5.6　对于连续函数 $f(t)$的梯形近似函数列 $f_n(t)$，及上面定义的 $I_n(\omega)$有

$$\| I_n \|^2 \xlongequal{\text{def}} EI_n^2 = E\left[\int_0^T f_n(t)\mathrm{d}B_t\right]^2 = \int_0^T |f_n(t)|^2 \mathrm{d}t.$$

于是，I_n 是 $L^2(\Omega)$中的 Cauchy 列. 从而 I_n 在 $L^2(\Omega)$中有极限，记为 $L^2(\Omega)\text{-}\lim\limits_{n\to+\infty} I_n$.

证明　注意两个梯形函数的差仍是梯形函数，故当 $n,m\to+\infty$时有

$$\begin{aligned} \| I_n - I_m \|^2 &= \int_0^T |f_n(t) - f_m(t)|^2 \mathrm{d}t \\ &\leqslant 2\left[\int_0^T |f_n(t) - f(t)|^2 \mathrm{d}t + \int_0^T |f(t) - f_m(t)|^2 \mathrm{d}t\right] \to 0. \end{aligned}$$

定义 5.3 对于连续函数 $f(t)$的梯形近似函数列 $f_n(t)$，及上面定义的 $I_n(\omega)$，定义

$$\int_0^T f(t)\mathrm{d}B_t \stackrel{\mathrm{def}}{=} L^2(\Omega)\text{-}\lim_{n\to+\infty} I(\omega) = L^2(\Omega)\text{-}\lim_{n\to+\infty}\int_0^T f_n(t)\mathrm{d}B_t.$$

即使在限制 Brown 运动的被积函数是确定性函数的情形，对于 $f(t)$的连续性要求仍然过强. 事实上，如果只要求 $f(t)$平方可积(它包括了大量不连续函数)，即只要求定积分 $\int_0^T f(t)^2\mathrm{d}t<+\infty$ (一般还可以推广为比定积分广的 Lebesgue 积分)，上面的路线仍是可行的. 具体地说，我们有下面的引理.

引理 5.7 如果 $\int_0^T f(t)^2\mathrm{d}t<+\infty$，那么存在梯形函数列 $f_n(t)$满足(也称为均方近似)

$$\int_0^T |f_n(t)-f(t)|^2\mathrm{d}t\to 0.$$

证明 这里只给出论证的纲要. 对于很小的 h，令

$$g_h(t)=\frac{1}{h}\int_t^{t+h} f(s)\mathrm{d}s$$

(事实上在 $t=T$ 处还需要作一点小小的修改，我们在此并不拘泥于过多的细节)，则 $g_h(t)$是连续函数，而且能"均方地近似"$f(t)$，即

$$\int_0^T |g_h(t)-f(t)|^2\mathrm{d}t\to 0.$$

再用梯形函数近似 $g_h(t)$(引理 5.5). 用取"对角线"方法就能得到 $f(t)$的均方近似梯形函数列.

定义 5.4 如果 $\int_0^T f(t)^2\mathrm{d}t<+\infty$，且梯形函数列 $f_n(t)$满足 $\int_0^T |f_n(t)-f(t)|^2\mathrm{d}t\to 0$ $\left(\text{这时}\int_0^T f_n(t)\mathrm{d}B_t \text{ 是 } L^2(\Omega)\text{中的 Cauchy 列}\right)$，定义

$$\int_0^T f(t)\mathrm{d}B_t = L^2(\Omega)\text{-}\lim_{n\to+\infty}\int_0^T f_n(t)\mathrm{d}B_t.$$

注意，在积分中的 t 是**哑变量**(dummy)，相当于"被求和"的变量，所以在结果中是不会出现的. 这个积分是一个随机变量. 不难看出，不同的均方近似函数序列的取法并不影响此积分.

完全类似地定义 $\int_a^b f(s)\mathrm{d}B_s$. 显然有

$$\int_a^b f(s)\mathrm{d}B_s = \int_0^b f(s)\mathrm{d}B_s - \int_0^a f(s)\mathrm{d}B_s.$$

将实值函数对 Brown 运动的不定限的积分视为随机过程：

$$\xi_t \stackrel{\mathrm{def}}{=} \int_0^t f(s)\mathrm{d}B_s.$$

那么，有实值函数对 Brown 运动的积分的性质.

命题 5.8 对于

$$\left\{\xi_t = \int_0^t f(u)\mathrm{d}B_u:\ t \geqslant 0\right\},$$

(1) 它是 Gauss 过程，其均值函数为 $m(t)=0$，相关函数为

$$E\left(\int_0^s f(u)\mathrm{d}B_u \int_0^t f(u)\mathrm{d}B_u\right) = \int_0^{s\wedge t} f(u)^2\mathrm{d}u.$$

(2) 它是独立增量过程. 而且 $\int_s^t f(u)\mathrm{d}B_u$ 与时刻 s 前的资料 $\{B_u:\ u\leqslant s\}$ 是独立的，即 $\int_s^t f(u)\mathrm{d}B_u$ 与任意 $\boldsymbol{B}_s$ 可知随机变量都是相互独立的.

(3) 如果 $|f(u)|\equiv 1(0\leqslant u)$，则 $\left\{\xi_t = \int_0^t f(u)\mathrm{d}B_u:\ t \geqslant 0\right\}$ 也是一个 Brown 运动.

(4) $E\mathrm{e}^{\int_0^t f(s)\mathrm{d}B_s} = \mathrm{e}^{\frac{1}{2}\int_0^t f(s)^2\mathrm{d}s}$.

证明 因为随机过程 $\left\{\xi_t^{(n)} \overset{\text{def}}{=\!=} \int_0^t f_n(u)\mathrm{d}B_u:\ t \geqslant 0\right\}$ 是均值为 0 的 Gauss 过程，且满足(2)，由定理 1.10 可知 $\left\{\xi_t = \int_0^t f(u)\mathrm{d}B_u:\ t \geqslant 0\right\}$ 也是均值为 0 的 Gauss 过程，并满足(2).

又若 $|f(u)|\equiv 1(0\leqslant u)$，那么

$$\xi_{t+s} - \xi_s = \int_s^{t+s} f(u)\mathrm{d}B_u \sim N\left(0,\int_s^{t+s} |f(u)|^2\mathrm{d}u\right) = N(0,s).$$

可知(3)成立. 又因为对于固定的 t，$\int_0^t f(u)\mathrm{d}B_u \sim N\left(0,\int_0^t f(u)^2\mathrm{d}u\right)$ 是 Gauss 随机变量，由 Gauss 随机变量的基本事实就得到(4).

命题 5.9

$$\mathrm{cov}\left(\int_a^b f(t)\mathrm{d}B_t,\int_a^b g(t)\mathrm{d}B_t\right) = E\left(\int_a^b f(t)\mathrm{d}B_t\int_a^b g(t)\mathrm{d}B_t\right) = \int_a^b f(t)g(t)\mathrm{d}t.$$

证明 我们有

$$\mathrm{var}\left(\int_a^b f(s)\mathrm{d}B_s\right) = E\left(\int_a^b f(s)\mathrm{d}B_s\right)^2 = \int_a^b f(s)^2\mathrm{d}s.$$

记

$$\xi = \int_a^b f(s)\mathrm{d}B_s,\quad \eta = \int_a^b g(u)\mathrm{d}B_u,$$

再利用

$$\mathrm{cov}(\xi,\eta) = \frac{1}{4}[\mathrm{var}(\xi+\eta) - \mathrm{var}(\xi-\eta)]$$

便得命题.

Ito 不定限积分是一个关于时间 t 的随机函数，也就是一个随机过程，它具有下面一

系列的鞅性质.

命题 5.10 $\left\{\xi_t = \int_0^t f(u)\mathrm{d}B_u : t \geqslant 0\right\}$ 是$(\mathbf{B}_t)$鞅.

证明 对于梯形函数 $f_n(t)$，对于 $s<t$，将 s 当作一个分点，加入到梯形函数 $f_n(t)$ 的梯形表达式中，再对 $\xi_t^{(n)} = \int_0^t f_n(u)\mathrm{d}B_u$ 利用 Brown 运动的独立增量性得到

$$
\begin{aligned}
E(\xi_t^{(n)} \mid \boldsymbol{B}_s) &= E(\xi_t^{(n)} - \xi_s^{(n)} \mid \boldsymbol{B}_s) + E(\xi_s^{(n)} \mid \boldsymbol{B}_s) \\
&= E(\xi_t^{(n)} - \xi_s^{(n)}) + \int_0^s f_n(u)\mathrm{d}B_u = \xi_s^{(n)}.
\end{aligned}
$$

所以 $\xi_t^{(n)} = \int_0^t f_n(s)\mathrm{d}B_s$ 是$(\mathbf{B}_t)$鞅. 再用推论 3.14，就得到 $\xi_t = \int_0^t f(s)\mathrm{d}B_s$ 是$(\mathbf{B}_t)$鞅.

命题 5.11 记

$$
\eta_t = \left(\int_0^t f(u)\mathrm{d}B_u\right)^2 - \int_0^t f(u)^2\mathrm{d}u,
$$

则$\{\eta_t : t\geqslant 0\}$是$(\mathbf{B}_t)$鞅，称为**随机积分的平方可积鞅**.

证明 对 $s<t$，因为 $\int_s^t f(u)\mathrm{d}B_u$ 与$\{B_u : u\leqslant s\}$是独立的，所以

$$
\begin{aligned}
E\left(\left[\int_0^s f(u)\mathrm{d}B_u\int_s^t f(u)\mathrm{d}B_u\right] \mid B_v : v\leqslant s\right) &= \left(\int_0^s f(u)\mathrm{d}B_u\right)E\left(\int_s^t f(u)\mathrm{d}B_u \mid B_v : v\leqslant s\right) \\
&= \left(\int_0^s f(u)\mathrm{d}B_u\right)E\left(\int_s^t f(u)\mathrm{d}B_u\right) = 0.
\end{aligned}
$$

从而

$$
\begin{aligned}
E\left(\left(\int_0^t f(u)\mathrm{d}B_u\right)^2 \mid B_u : u\leqslant s\right) &= E\left(\left[\int_0^s f(u)\mathrm{d}B_u + \int_s^t f(u)\mathrm{d}B_u\right]^2 \mid B_u : u\leqslant s\right) \\
&= \left(\int_0^s f(u)\mathrm{d}B_u\right)^2 + E\left(\int_s^t f(u)\mathrm{d}B_u\right)^2 \\
&= \left(\int_0^s f(u)\mathrm{d}B_u\right)^2 + \int_s^t f^2(u)\mathrm{d}u.
\end{aligned}
$$

于是

$$
\begin{aligned}
E(\eta_t \mid B_u : u\leqslant s) &= E\left[\left(\int_0^t f(u)\mathrm{d}B_u\right)^2 - \int_0^t f(u)^2\mathrm{d}u \mid B_u : u\leqslant s\right] \\
&= \left(\int_0^s f(u)\mathrm{d}B_u\right)^2 + \int_s^t f^2(u)\mathrm{d}u - \int_0^t f(u)^2\mathrm{d}u = \eta_s.
\end{aligned}
$$

命题 5.12 记

$$
\zeta_t = \mathrm{e}^{\int_0^t f(s)\mathrm{d}B_s - \frac{1}{2}\int_0^t f(s)^2\mathrm{d}s},
$$

则$\{\zeta_t : t\geqslant 0\}$也是$(\mathbf{B}_t)$鞅，称为**随机积分的指数鞅**.

证明 因为 $\int_s^t f(s)\mathrm{d}B_s - \dfrac{1}{2}\int_s^t f(s)^2\mathrm{d}s$ 与$\{B_u : u\leqslant s\}$独立，且

$$\int_s^t f(u)\mathrm{d}B_u \sim N\left(0,\int_s^t |f(u)|^2\mathrm{d}u\right),$$

所以由命题 5.8 的(4)，有

$$E(\mathrm{e}^{\int_s^t f(s)\mathrm{d}B_s-\frac{1}{2}\int_s^t f(s)^2\mathrm{d}s} \mid B_u: u\leqslant s) = E(\mathrm{e}^{\int_s^t f(s)\mathrm{d}B_s-\frac{1}{2}\int_s^t f(s)^2\mathrm{d}s}) = \mathrm{e}^{-\frac{1}{2}\int_s^t f(s)^2\mathrm{d}s}E\,\mathrm{e}^{\int_s^t f(s)\mathrm{d}B_s} = 1.$$

从而

$$E(\zeta_t \mid B_u: u\leqslant s) = \zeta_s E(\mathrm{e}^{\int_s^t f(s)\mathrm{d}B_s-\frac{1}{2}\int_s^t f(s)^2\mathrm{d}s} \mid B_u: u\leqslant s) = \zeta_s.$$

注　若 $f(t)$连续可微，则可以将随机积分化成为如下的按轨道的普通积分：

$$\int_a^b f(t)\mathrm{d}B_t = [f(b)B_b - f(a)B_a] - \int_a^b f'(t)B_t\mathrm{d}t.$$

为此只需验证两边的差的方差为 0.

特别地有

$$\int_0^t s\mathrm{d}B_s = tB_t - \int_0^t B_s\mathrm{d}s,$$

即

$$\int_0^t B_u\mathrm{d}u = tB_t - \int_0^t s\mathrm{d}B_s.$$

它建立了积分 Brown 运动与对于 Brown 运动的积分间的联系.

实值函数对 Brown 运动的积分在物理中的应用过于受限制. 而随机积分的引进，主要来自物理等领域，在那里研究的运动，是一个受到随机干扰的系统的输出，最简单的情形是输入受到了像 Brown 运动那样的随机力的干扰的一个常微分方程的解，即随机微分方程的解. 由于输入是 Brown 运动的随机干扰，方程的解中所涉及的积分就是对 Brown 运动的积分. 又由于一般地随机干扰的系数不仅依赖于时间，而且更常见的是还依赖于位置，故而在经过随机干扰后的系数就是随机过程. 因此，所涉及的积分的被积函数一般地将是随机过程. 这样，我们就必须定义随机过程对 Brown 运动的积分.

5.2.2　随机函数(随机过程)对 Brown 运动的积分

随机函数(随机过程)对 Brown 运动的积分是随机微积分的主体.

一般的随机过程对 Brown 运动的积分，称为 **Skorohod 积分**，这种积分的定义较为复杂，但是对于被积随机过程的限制较少，而这种积分的性质也较为特殊，故而目前其应用范围还相对地受局限，更多地见之于较为理论性的研究中，例如，物理中的规范场的数学理论基础，或随机粒子对于随机介质边界的反射运动或穿透运动等的研究中. 本书不予探究.

但是如前面所论述的，受到 Brown 运动$\{B_t: t\geqslant 0\}$所随机干扰的物理或工程系统，其输出的解中所涉及的积分的被积随机过程，从其受到的干扰的历史看，应该是(B_t)可知的(即适应的)随机过程. 这说明我们需要的积分是(B_t)可知的随机过程对(B_t)的积分. 这就是 20 世纪中叶日本数学家 Ito 所定义的随机积分. 其后，人们称这种积分为 **Ito 积分**. 也

就是说，Ito 积分源自 Brown 运动驱动的随机微分方程的解的表达与分析.

让我们回忆，当基本事件 ω 固定时，随机过程 $\Phi_t \overset{\text{def}}{=\!=} \Phi_t(\omega)$ 的一个样本函数，或轨道，就是将 $\Phi_t(\omega)$ 作为 t 的函数. 如果样本函数都是 t 的连续函数，则称随机过程 Φ_t 具有连续轨道，或称 Φ_t 为**连续的随机过程**. 而 Brown 运动就是连续的随机过程.

(1) $(\boldsymbol{B}_t)$ 可知的随机过程对 Brown 运动 $\{\boldsymbol{B}_t: t\geqslant 0\}$ 的 Ito 积分

考虑区间 $[a,b]$ 的划分：

$$a = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n}^{(n)} = b, \quad \lambda_n \overset{\text{def}}{=\!=} \max_{0\leqslant k\leqslant N_n-1} \{t_{k+1}^{(n)} - t_k^{(n)}\} \to 0,$$

并定义 (B_t) 可知的随机过程 Φ_t 的如下积分和

$$I_n(\omega) \overset{\text{def}}{=\!=} \sum_{k=0}^{N_n-1} \Phi_{t_k^{(n)}}(\omega)(B_{t_{k+1}^{(n)}}(\omega) - B_{t_k^{(n)}}(\omega)). \tag{5.1}$$

如同被积函数为实值函数的情形，将随机变量 $I_n \overset{\text{def}}{=\!=} I_n(\omega)\,(\omega\in\Omega)$ 看成 $L^2(\Omega)$ 中的元. 由于数学上纯技术的原因，现在的相应于定义 5.2 的随机版本，应作如下的修改.

定义 5.2′　如果 $(0,T]$ 的子区间 $(s_i,t_i]\,(i=1,2,\cdots,m)$ 两两不相交，且随机过程 Φ_t 是梯形随机过程，即它有表达式

$$\Phi_t = \sum_{i=1}^{m} \Phi_{s_i} I_{(s_i,t_i]}(t), \tag{5.2}$$

且其中的系数随机变量 Φ_{s_i} 为 $\boldsymbol{B}_{s_i} \overset{\text{def}}{=\!=} \{B_u: u\leqslant s_i\}$ 可知的，则称随机过程 Φ_t 为 (B_t) **可料的梯形过程**.

一般的非梯形随机过程的 (B_t) 可料性概念是随机过程的一般理论中的一个深入的概念，它远比随机过程的 (B_t) 可知性概念复杂，要求的条件也强得多，我们不准备介绍. 在这里我们只给出梯形随机过程的 (B_t) 可料性概念. 一个梯形随机过程的 (B_t) 可料性，要求随机变量 Φ_{t_i} 是 $\boldsymbol{B}_{s_i}$ 可知的，而一个梯形随机过程的 (B_t) 可知性，则只要求随机变量 Φ_{t_i} 是 $\boldsymbol{B}_{t_i}$ 可知的，又因为 $s_i<t_i$，$\boldsymbol{B}_{s_i}$ 代表 s_i 前历史，它比代表 t_i 前历史的 $\boldsymbol{B}_{t_i}$ 少得多，所以 (B_t) 可料性要求比 (B_t) 可知性要求高得多.

对于任意固定的基本事件 ω，定义 (B_t) 可料的梯形过程 Φ_t 对 Brown 运动的积分为

$$\int_0^T \Big[\sum_{i=0}^{m} \Phi_{s_i} I_{(s_i,t_i]}(t)\Big] \mathrm{d}B_t(\omega) \overset{\text{def}}{=\!=} \sum_{i=0}^{m} \Phi_{s_i}(B_{t_i} - B_{s_i}), \tag{5.3}$$

称为可料的梯形过程 Φ_t 的 **Ito 积分**. 显然，它是一个随机变量.

同样，一个 (B_t) 可料梯形过程的不同表示在不计零概率事件差异下不会影响积分. (B_t) 可料的梯形过程 Φ_t 的 Ito 积分 $\int_0^T \Phi_t \mathrm{d}B_t(\omega)$ 通常简单地记为 $\int_0^T \Phi_t \mathrm{d}B_t$.

引理 5.4′　(B_t) 可料的梯形过程 Φ_t 的 Ito 积分满足

$$E\Big[\int_0^T \Phi_t \mathrm{d}B_t\Big] = 0,$$

$$\left\|\int_0^T \Phi_t \mathrm{d}B_t\right\|^2 = \sum_{i=0}^m \|\Phi_{s_i}\|^2(t_i - s_i) = \int_0^T \|\Phi_t\|^2 \mathrm{d}t. \tag{5.4}$$

后者是随机梯形函数 $\|\Phi_t\|^2(\overset{\text{def}}{=\!=} E|\Phi_t|^2)$的定积分.

证明　利用 Φ_t 的(B_t)可料性，以及条件期望的性质，我们得到

$$\begin{aligned}
E\left(\int_0^T \Phi_t \mathrm{d}B_t\right)^2 &= E\left[\sum_{i=0}^m \Phi_{s_i}(B_{t_i} - B_{s_i})\right]^2 \\
&= E\left[\sum_{i=0}^m \Phi_{s_i}^2(B_{t_i} - B_{s_i})^2 + 2\sum_{j<k}\Phi_{s_j}\Phi_{s_k}(B_{t_j} - B_{s_j})(B_{t_k} - B_{s_k})\right] \\
&= E\left[\sum_{i=0}^m E(\Phi_{s_i}^2(B_{t_i} - B_{s_i})^2 \mid \boldsymbol{B}_{s_i}) + 2\sum_{j<k}E(\Phi_{s_j}\Phi_{s_k}(B_{t_j} - B_{s_j})(B_{t_k} - B_{s_k}) \mid \boldsymbol{B}_{s_k})\right] \\
&= E\left[\sum_{i=0}^m \Phi_{s_i}^2 E((B_{t_i} - B_{s_i})^2 \mid \boldsymbol{B}_{s_i}) + 2\sum_{j<k}\Phi_{s_j}(B_{t_j} - B_{s_j})E(\Phi_{s_k}(B_{t_k} - B_{s_k}) \mid \boldsymbol{B}_{s_k})\right] \\
&= E\left[\sum_{i=0}^m \Phi_{s_i}^2 E(B_{t_i} - B_{s_i})^2 + 2\sum_{j<k}\Phi_{s_j}(B_{t_j} - B_{s_j})E(\Phi_{s_k}((B_{t_k} - B_{s_k}) \mid \boldsymbol{B}_{s_k})\boldsymbol{B}_{s_k})\right] \\
&= E\left[\sum_{i=0}^m \Phi_{s_i}^2(t_i - s_i)\right] = \sum_{i=0}^m E\Phi_{s_i}^2(t_i - s_i) = \sum_{i=0}^m \|\Phi_{s_i}\|^2(t_i - s_i).
\end{aligned}$$

这个引理的证明. 充分利用了条件期望的以下性质，即对于信息源为可知的随机变量，在对信息源取条件期望时，可以当作“常数”一样处理，而梯形随机过程的可料性要求，不仅保证了在区间$(t_k, t_{k+1}]$上随机变量 Φ_t 是$\{B_s: s\leqslant t\}$可知的，而且要求它对于更小的信息源$\{B_s: s\leqslant t_k\}$也是可知的，正是这个加强的要求，确保了在对$\{B_s: s\leqslant t_k\}$求条件期望时，其系数随机变量可以像常数一样从条件期望运算符号中拿出去. 故而梯形随机过程(B_t)可料性的使用是 Ito 定义随机积分时的技术突破点.

令

$$\mathcal{L}^2 \overset{\text{def}}{=\!=} \left\{(B_t)\text{ 可知的随机过程 }\Phi_t\text{，且}\int_0^T E\mid\Phi_t(\omega)\mid^2\mathrm{d}t < +\infty\right\}. \tag{5.5}$$

对$\mathcal{L}^2$中的元素 $\Phi \overset{\text{def}}{=\!=} \{\Phi_t: t\geqslant 0\}$，$\Psi \overset{\text{def}}{=\!=} \{\Psi_t: t\geqslant 0\}$，可以定义如下的内积：

$$((\Phi, \Psi)) \overset{\text{def}}{=\!=} \int_0^T E(\Phi_t \Psi_t)\mathrm{d}t. \tag{5.6}$$

（在 Φ_t，Ψ_t 都是连续的随机过程时，右方的积分是普通的定积分，在一般的情形，右方的积分是定积分的一种推广，在实变函数理论中称为 **Lebesgue 积分**. 当被积函数的定积分存在时，这种积分与定积分是相等的，而当被积函数的连续程度不好而使定积分不存在时，Lebesgue 积分常会存在. 然而 Lebesgue 积分很难直接计算，因此它更多地用在理论分析，通过它建立某些理论上的重要联系）和范数

$$|||\,\Phi\,|||^2 \overset{\text{def}}{=\!=} \int_0^T E\mid\Phi_t\mid^2\mathrm{d}t. \tag{5.7}$$

不难验证$\mathcal{L}^2$具有与$L^2(\Omega)$非常类似的性质,即在内积$((\Phi,\Psi))$下成为 Hilbert 空间.

对于$\mathcal{L}^2$中的连续随机过程Φ_t,及区间$[0,T]$的任意划分:

$$0=t_0^{(n)}<t_1^{(n)}<\cdots<t_{N_n}^{(n)}=T,\quad \lambda_n \overset{\text{def}}{=\!=} \max_{0\leqslant k\leqslant N_n-1}\{t_{k+1}^{(n)}-t_k^{(n)}\}\to 0,$$

对于固定的n,可以直接验证,和数

$$\Phi_t^{(n)} \overset{\text{def}}{=\!=} \sum_k \Phi_{t_k^{(n)}} I_{(t_k^{(n)}\wedge t,t_{k+1}^{(n)}\wedge t)}(t)+\Phi_0 I_{\{0\}}(t)$$

(注意:取的是左端点处的值$\Phi_{t_k}^{(n)}$!)是一个(B_t)可料的梯形随机过程,称为近似Φ_t的(B_t)可料的梯形随机过程列.

引理 5.5′ 对于$\mathcal{L}^2$中的连续随机过程Φ_t,及其近似的(B_t)可料的梯形随机过程列$\Phi_t^{(n)}$有

$$|||\ \Phi^{(n)}-\Phi\ |||^2=E\int_0^T|\Phi_t^{(n)}-\Phi_t|^2\mathrm{d}t\to 0. \tag{5.8}$$

更一般地,我们还有下面的引理.

引理 5.5″ 对于$\mathcal{L}^2$中的随机过程Φ_t,必定存在(B_t)可料的梯形随机过程列$\Phi_t^{(n)}$,使得在$n\to+\infty$时有

$$|||\ \Phi^{(n)}-\Phi\ |||^2=E\int_0^T|\Phi_t^{(n)}-\Phi_t|^2\mathrm{d}t\to 0, \tag{5.8′}$$

从而$\Phi^{(n)}$是$\mathcal{L}^2$中的 Cauchy 序列.

易见引理 5.5′是引理 5.5″的特殊情形.引理 5.5″也是定义 Ito 积分的技术突破点.引理 5.5″的证明需要实变函数论中的逼近技巧,本书从略.

推论 5.6′ 对于引理 5.5″中的随机过程Φ_t,及其近似的(B_t)可料的梯形随机过程列$\Phi_t^{(n)}$有

$$\left\|\int_0^T\Phi_t^{(n)}\mathrm{d}B_t\right\|^2=\int_0^T\|\Phi_t^{(n)}\|^2\mathrm{d}t=|||\ \Phi^{(n)}\ |||^2, \tag{5.9}$$

且$\int_0^T\Phi_t^{(n)}\mathrm{d}B_t$是$L^2(\Omega)$中的 Cauchy 列.

证明与推论 5.6 类似.

定义 5.4′ 对于$\mathcal{L}^2$中的随机过程Φ_t,且$\Phi_t^{(n)}$是(B_t)可料的梯形随机过程列,满足

$$|||\ \Phi^{(n)}-\Phi\ |||^2=E\int_0^T|\Phi_t^{(n)}-\Phi_t|^2\mathrm{d}t\to 0. \tag{5.9′}$$

$\left(\text{这时}\int_0^T\Phi_t^{(n)}\mathrm{d}B_t\text{ 是 }L^2\text{ 中的 Cauchy 列}\right)$定义

$$\int_0^T\Phi_t\mathrm{d}B_t=L^2(\Omega)\text{-}\lim_{n\to+\infty}\int_0^T\Phi_t^{(n)}\mathrm{d}B_t, \tag{5.10}$$

称为 Ito 积分.易见,使用不同的(B_t)可料梯形随机过程近似列,在不计零概率事件差异下不会影响积分的定义.同样$\int_0^T\Phi_t\mathrm{d}B_t$常常简记为$\int_0^T\Phi\mathrm{d}B$.

类似地定义 $\int_a^b \Phi_t \mathrm{d}B_t$. 则仍有

$$\int_a^b \Phi_t \mathrm{d}B_t = \int_0^b \Phi_t \mathrm{d}B_t - \int_0^a \Phi_t \mathrm{d}B_t.$$

注　我们再一次强调，对于$\mathcal{L}^2$中的连续随机过程 Φ_t，其近似梯形过程列 $\Phi^{(n)}$ 的(B_t)可料性，要求相应于在每个小区间$(t_k^{(n)}, t_{k+1}^{(n)}]$上，限定取 $\Phi_{t_k^{(n)}}$ 作为 Φ 在此小区间上的近似. 而不能像在普通函数的定积分中那样，可以取 Φ 在此小区间上任意的一点处的值.

注意，对 Brown 运动在小区间上取不同的点得到的不同的随机和数，在取极限时有明显的差别，是不能再忽视的. 事实上，这里相应于普通定积分中的自变量的差分的项是：$\Delta B_{t_k^{(n)}} = B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}}$，它们是随机变量. 虽然由于 Brown 运动是连续的随机过程有 $\lim\limits_{n\to+\infty} \Delta B_k^{(n)}(\omega)=0$，但是，对于不同的 ω，它们趋于 0 的速度是很不一致的. 粗略地说（确切的说法可见下面的引理 5.13），平均地有

$$(\Delta B_k^{(n)})^2 \approx E(\Delta B_k^{(n)})^2 = \Delta t_k^{(n)} \quad (\text{或者说，}(\mathrm{d}B_t)^2 = \mathrm{d}t),$$

也就是说，平均地 $\Delta B_k^{(n)}$ 趋于 0 的速度为 $\sqrt{\Delta t_k^{(n)}}$，它大大地慢于 $\Delta t_k^{(n)}$. 从而在区间$(t_k^{(n)}, t_{k+1}^{(n)}]$上取不同的点处的值作为 Φ_t 的近似，所得的近似和之间的差别是不可忽略的.

引理 5.13　对于区间$[0,T]$上的划分，记

$$\Delta B_{t_k^{(n)}} = B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}},$$

则当

$$\lambda_n = \max_{0\leqslant k\leqslant N_n-1} \{t_{k+1}^{(n)} - t_k^{(n)}\} \to 0$$

时有

$$E\left|\sum_k (\Delta B_{t_k^{(n)}})^2 - E\left[\sum_k (\Delta B_{t_k^{(n)}})^2\right]\right|^2 = E\left|\sum_k (\Delta B_{t_k^{(n)}})^2 - T\right|^2 \to 0. \quad (5.11)$$

证明　由 Brown 运动的性质

$$\Delta B_{t_k^{(n)}} = B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}} \sim N(0, \Delta t_k^{(n)}),$$

有

$$E\sum_k (\Delta B_{t_k^{(n)}})^2 = t, \quad E(\Delta B_{t_k^{(n)}})^4 = 3(\Delta t_k^{(n)})^2.$$

而 Brown 运动的独立增量性，保证了对于 $i\neq j$，还有

$$E([(\Delta B_{t_i^{(n)}})^2 - \Delta t_i^{(n)}][(\Delta B_{t_j^{(n)}})^2 - \Delta t_j^{(n)}]) = 0.$$

于是

$$\begin{aligned} E\left|\sum_k (\Delta B_{t_k^{(n)}})^2 - t\right|^2 &= E\left(\sum_k [(\Delta B_{t_k^{(n)}})^2 - \Delta t_k^{(n)}]\right)^2 \\ &= E\left(\sum_k [(\Delta B_{t_k^{(n)}})^2 - \Delta t_k^{(n)}]^2\right) \\ &\quad + 2E\left(\sum_{i\neq j} [(\Delta B_{t_i^{(n)}})^2 - \Delta t_i^{(n)}][(\Delta B_{t_j^{(n)}})^2 - \Delta t_j^{(n)}]\right) \end{aligned}$$

$$= \sum_k E[(\Delta B_{t_k^{(n)}})^2 - \Delta t_k^{(n)}]^2$$
$$= \sum_k E[(\Delta B_{t_k^{(n)}})^4 + (\Delta t_k^{(n)})^2 - 2(\Delta B_{t_k^{(n)}})^2(\Delta t_k^{(n)})]$$
$$= 2\sum_k (\Delta t_k^{(n)})^2 \leqslant 2\lambda_n \sum_k \Delta t_k^{(n)} = 2T\lambda_n \to 0 \quad (n \to +\infty).$$

例 5.1 我们有

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t^2 - \frac{t}{2}. \tag{5.12}$$

这个结果与普通微积分的不定积分公式的不同之处是这里多了一项$-\frac{1}{2}t$.

证明

$$I_n \overset{\text{def}}{=\!=} \sum_k B_{t_k^{(n)}}(B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})$$
$$= \frac{1}{2}\sum_k (2B_{t_k^{(n)}}B_{t_{k+1}^{(n)}} - 2B_{t_k^{(n)}}^2) = \frac{1}{2}\sum_k [\Delta(B_{t_k^{(n)}})^2 - (\Delta B_{t_k^{(n)}})^2]$$
$$= \frac{1}{2}B_t^2 - \frac{1}{2}\sum_k (\Delta B_{t_k^{(n)}})^2.$$

于是由引理 5.13 推出

$$E\left| I_n - \left(\frac{1}{2}B_t^2 - \frac{t}{2}\right)\right|^2 \to 0.$$

即 I_n 在 $L^2(\Omega)$中收敛到$\frac{1}{2}(B_t^2 - t)$$\left(\text{而不是}\frac{1}{2}B_t^2!\right)$.

引理 5.14 对于有界连续函数 f 有

$$E\left|\sum_k f(B_{t_k^{(n)}})(\Delta B_{t_k^{(n)}})^2 - \int_0^T f(B_t)\mathrm{d}t\right|^2 \to 0. \tag{5.13}$$

证明 记 $M = \max\limits_x |f(x)|$,于是

$$E\left|\sum_k f(B_{t_k^{(n)}})(\Delta B_{t_k^{(n)}})^2 - \int_0^T f(B_t)\mathrm{d}t\right|^2$$
$$= E\left|\sum_k f(B_{t_k^{(n)}})[(\Delta B_{t_k^{(n)}})^2 - \Delta t_k^{(n)}] + \left[\sum_k f(B_{t_{k-1}^{(n)}})\Delta t_k^{(n)} - \int_0^T f(B_t)\mathrm{d}t\right]\right|^2.$$
$$\leqslant 2ME\left|\sum_k [(\Delta B_{t_k^{(n)}})^2 - \Delta t_k^{(n)}]\right|^2 + 2E\left[\sum_k f(B_{t_{k-1}^{(n)}})\Delta t_k^{(n)} - \int_0^T f(B_t)\mathrm{d}t\right]^2$$
$$\leqslant 2ME\left|\sum_k [(\Delta B_{t_k^{(n)}})^2 - \Delta t_k^{(n)}]\right|^2 + 2E\left[\sum_k f(B_{t_{k-1}^{(n)}})\Delta t_k^{(n)} - \int_0^T f(B_t)\mathrm{d}t\right]^2.$$

利用引理 5.13 推出,当 $\lambda_n \to 0$ 时,上面的和数中的第一项趋于 0. 又因为对于一切 ω 都有 $\sum_k f(B_{t_{k-1}^{(n)}})\Delta t_k^{(n)} \to \int_0^T f(B_t)\mathrm{d}t$,故而第二项也趋于 0(这里利用了在期望号内取极限,可

以这样做的严格论证需用测度论中的控制收敛定理. 此外，函数 f 的有界性假定也可以去除，并不影响引理的结论).

例 5.2　如果我们在例 5.1 的积分和中，如果用 $B_{t_{k+1}^{(n)}}$ 代替 $B_{t_k^{(n)}}$，并记

$$H_n = \sum_k B_{t_{k+1}^{(n)}} \Delta B_{t_k^{(n)}},$$

那么有

$$H_n = \Phi^{(n)} + \sum_k (\Delta B_{t_k^{(n)}})^2.$$

于是

$$E\left|H_n - \left(\frac{1}{2}B_t^2 + \frac{t}{2}\right)\right|^2 \to 0.$$

注　结合例 5.1 与例 5.2，我们得到

$$E\left|\sum_k \frac{B_{t_k^{(n)}} + B_{t_{k+1}^{(n)}}}{2}\Delta B_{t_k^{(n)}} - \frac{1}{2}B_t^2\right|^2 \to 0. \tag{5.14}$$

定义 5.5(Stratonovich 随机积分)　若 $f(t,x)$为连续可微实值函数，则

$$\sum_k f\left(t_{t_k^{(n)}}, \frac{B_{t_k^{(n)}} + B_{t_{k+1}^{(n)}}}{2}\right)\Delta B_{t_k^{(n)}}$$

在 $L^2(\Omega)$中的极限，称为 **Stratonovich 积分**，记为

$$\int_0^T f(t,B_t) \circ \mathrm{d}B_t.$$

又若 ξ_t 是一个(B_t)可知的连续随机过程，那么，类似地可以定义

$$\int_0^T f(t,\xi_t) \circ \mathrm{d}B_t \overset{\text{def}}{=} L^2(\Omega)\text{-}\lim_{n\to+\infty}\sum_k f\left(t_{t_k^{(n)}}, \frac{\xi_{t_k^{(n)}} + \xi_{t_{k+1}^{(n)}}}{2}\right)\Delta B_{t_k^{(n)}}.$$

用与例 5.2 类似的推理，可以得到 Stratonovich 积分与 Ito 积分的如下基本关系：

$$\int_0^t f(B_s) \circ \mathrm{d}B_s = \int_0^t f(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f'(B_s)\mathrm{d}s. \tag{5.15}$$

更一般地有

$$\int_0^T f(t,B_t) \circ \mathrm{d}B_t = \int_0^T f(t,B_t)\mathrm{d}B_t + \frac{1}{2}\int_0^T \frac{\partial f}{\partial x}(t,B_t)\mathrm{d}t. \tag{5.15$'$}$$

这说明 Stratonovich 积分总可以化为 Ito 积分，反之则不然.

注意若 ξ_t 是一个(B_t)可知的连续随机过程，那么，一般地

$$\int_0^t f(\xi_s) \circ \mathrm{d}B_s \neq \int_0^t f(\xi_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f'(\xi_s)\mathrm{d}s. \tag{5.15$''$}$$

其确切形式依赖于 ξ_t，而其特殊情形可以由后面的 Ito 公式得到.

例 5.3

$$\int_0^t B_t \circ \mathrm{d}B_t = \frac{1}{2}B_t^2.$$

这是例 5.2 的注的复述.

例 5.4 如果函数 f 有原函数 F,则利用积分和取极限可以直接验证

$$\int_0^t f(B_t) \circ \mathrm{d}B_t = F(B_t) - F(B_0).$$

这里 Stratonovich 积分的结果完全与普通微积分中的结果一致. 而 Ito 积分就没有这样简单的形式. 这是 Stratonovich 积分的优点. 但是,Ito 随机积分更便于作技术处理. 下面我们将说明,不定上限的 Ito 积分作为随机过程是鞅(对于后文 5.3.2 节中推广的被积随机过程的情形,实际上所对应的是所谓局部鞅). 这说明 Ito 积分便于利用鞅论这一有力工具. 此外,下文我们将说明,对于 Brown 运动的差商的光滑逼近,相应的极限是 Stratonovich 积分,而不是 Ito 积分. 所以,在很多情况下,数学家更喜欢使用 Ito 积分,因为他们要对随机积分的许多结论给出理论推导. 而物理学家则更喜欢使用 Stratonovich 积分. 因为后者的运算规律与普通微积分相似.

下面讨论 Ito 积分的基本性质.

首先是类似于普通积分的性质.

(1) 线性性质

$$\int_0^t (\Phi + \Psi)\mathrm{d}B = \int_0^t \Phi \mathrm{d}B + \int_0^t \Psi \mathrm{d}B,$$

$$\int_0^t c\Phi \mathrm{d}B = c\int_0^t \Phi \mathrm{d}B.$$

对 $\{B_u: u \leqslant a\}$ 可知的有界随机时间 η,还有

$$\int_a^b \eta\Phi \mathrm{d}B = \eta\int_a^b \Phi \mathrm{d}B.$$

(2) 可加性:对 $a<b<c$ 有

$$\int_a^c \Phi \mathrm{d}B = \int_a^b \Phi \mathrm{d}B + \int_b^c \Phi \mathrm{d}B.$$

(3) 对任意有界的 (B_t) 停时 τ,假定 $\tau \leqslant T$,则可以定义

$$\int_0^\tau \Phi \mathrm{d}B \overset{\mathrm{def}}{=\!=} \int_0^T \Phi_t I_{(0,\tau]}(t)\mathrm{d}B_t.$$

不定限的 Ito 积分 $\int_0^t \Phi \mathrm{d}B$ 作为随机过程不再是 Gauss 过程,而且满足

(4) $E\left(\int_0^t \Phi_s \mathrm{d}B_s\right) = 0, E\left(\int_u^t \Phi_s \mathrm{d}B_s \int_u^t \Psi_s \mathrm{d}B_s\right) = \int_u^t E(\Phi_s \Psi_s)\mathrm{d}s.$

(性质(4)是如下的引理 5.15 的推论).

引理 5.15 对于 $\mathcal{L}^2$ 中的随机过程 Φ_t,及 $s<t$ 有

$$E\left(\left[\int_s^t \Phi_u \mathrm{d}B_u\right]^2 \,\Big|\, \boldsymbol{B}_s\right) = E\left(\int_s^t \Phi_u^2 \mathrm{d}u \mid \boldsymbol{B}_s\right).$$

证明 我们先假定 Φ_t 是 (B_t) 可料的梯形过程. 将 s,t 作为两个分点增加到划分中

后，对于 $s<t_k<t_j$，由 Φ_t 的 (B_t) 可料性推出

$$E(\Phi_{t_k}^2(\Delta B_{t_k})^2 \mid \boldsymbol{B}_s)=E[E(\Phi_{t_k}^2(\Delta B_{t_k})^2 \mid \boldsymbol{B}_{t_k}) \mid \boldsymbol{B}_s]$$
$$=E[\Phi_{t_k}^2 E((\Delta B_{t_k})^2 \mid \boldsymbol{B}_{t_k}) \mid \boldsymbol{B}_s]=E[\Phi_{t_k}^2 \Delta t_k \mid \boldsymbol{B}_s].$$

另一方面，由于 Brown 运动是鞅，我们知道 $E(\Delta B_{t_j} \mid \boldsymbol{B}_{t_j})=0$，故而

$$E(\Phi_{t_k}\Delta B_{t_k}\Phi_{t_j}\Delta B_{t_j} \mid \boldsymbol{B}_s)=E[E(\Phi_{t_k}\Delta B_{t_k}\Phi_{t_j}\Delta B_{t_j} \mid \boldsymbol{B}_{t_k}) \mid \boldsymbol{B}_s]$$
$$=E[\Phi_{t_k}\Phi_{t_j}\Delta B_{t_k}E(\Delta B_{t_j} \mid \boldsymbol{B}_{t_j}) \mid \boldsymbol{B}_s]=0.$$

于是

$$E\left(\left(\int_s^t \Phi_u \mathrm{d}B_u\right)^2 \mid \boldsymbol{B}_s\right)=E\left(\left(\sum_{t_k>s}\Phi_{t_k}\Delta B_{t_k}\right)^2 \mid \boldsymbol{B}_s\right)$$
$$=E\left(\left(\sum_{t_k>s}\Phi_{t_k}^2(\Delta B_{t_k})^2+2\sum_{k<j,t_k>s}\Phi_{t_k}\Delta B_{t_k}\Phi_{t_j}\Delta B_{t_j}\right) \mid \boldsymbol{B}_s\right)$$
$$=E\left[\sum_{t_k>s}\Phi_{t_k}^2\Delta t_k \mid \boldsymbol{B}_s\right]=E\left(\int_s^t \Phi_u^2 \mathrm{d}u \mid \boldsymbol{B}_s\right).$$

所以，引理在 Φ_t 是 (B_t) 可料的梯形过程是正确的.

对于 $\mathcal{L}^2$ 中的 Φ_t，由引理 5.5′，可以用 (B_t) 可料的梯形过程列近似，用推论 3.10 就得到这个引理的证明.

定理 5.16　对于 $\mathcal{L}^2$ 中的随机过程 Φ_t，在 t 变化时，有

(A) $\left\{\xi_t \overset{\text{def}}{=\!=} \int_0^t \Phi_s \mathrm{d}B_s\right\}$ 是 (B_t) 鞅　(这是 Brown 运动鞅的推广).　　(5.16)

(B) $\left\{\eta_t \overset{\text{def}}{=\!=} \left(\int_0^t \Phi_s \mathrm{d}B_s\right)^2-\int_0^t \Phi_s^2 \mathrm{d}s\right\}$ 是 (B_t) 鞅(这是 Brown 运动平方可积鞅的推广).

(5.17)

(C) 若 Φ_t 是有界的随机过程，即 $|\Phi_t|<$ 某个常数，则

$$\left\{\zeta_t \overset{\text{def}}{=\!=} \mathrm{e}^{\int_0^t \Phi_s \mathrm{d}B_s-\frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s}\right\} \tag{5.18}$$

是 (B_t) 鞅(这是 Brown 运动指数鞅的推广). 并且在此条件下，进而对于任意 $\varepsilon>0$，随机积分的最大值有“概率估计”：

$$P\left(\max_{t\leqslant 1}\left\{\int_0^t \Phi_s \mathrm{d}B_s-\frac{\alpha}{2}\int_0^t \Phi_s^2 \mathrm{d}s\right\}>\varepsilon\right)\leqslant \mathrm{e}^{-\alpha\varepsilon}. \tag{5.19}$$

证明　(A) 的证明首先假定 Φ_t 是 (B_t) 可料的梯形随机过程，故对于

$$\boldsymbol{B}_s=\{B_u:u\leqslant s\}$$

随机变量 $\int_0^s \Phi_u \mathrm{d}B_u$ 是 $\boldsymbol{B}_s$ 可知的，从而

$$E\left(\int_0^s \Phi_u \mathrm{d}B_u \middle| \boldsymbol{B}_s\right)=\int_0^s \Phi_u \mathrm{d}B_u.$$

再则，当 $s<t$ 时，将 s,t 作为两个分点增加到划分中后，对于 $s\leqslant t_k<t_{k+1}$，由 Φ_t 的 (B_t) 可

料性我们推出

$$E(\Phi_{t_k}\Delta B_{t_k}\mid \boldsymbol{B}_s)=E[E(\Phi_{t_k}\Delta B_{t_k}\mid \boldsymbol{B}_{t_k})\mid \boldsymbol{B}_s]=E[\Phi_{t_k}E(\Delta B_{t_k}\mid \boldsymbol{B}_{t_k})\mid \boldsymbol{B}_s]=0.$$

对于 k 求和便得

$$E\left(\int_s^t \Phi_u \mathrm{d}B_u \,\middle|\, \boldsymbol{B}_s\right)=0.$$

合起来就是

$$E\left(\int_0^t \Phi_u \mathrm{d}B_u \,\middle|\, \boldsymbol{B}_s\right)=\int_0^s \Phi_u \mathrm{d}B_u.$$

其次,对于$\mathcal{L}^2$中的随机过程 Φ_t,它可以用(B_t)可料的梯形随机过程 $\Phi_t^{(n)}$ 近似. 而上面我们已经证明了 $\int_0^t \Phi_s^{(n)}\mathrm{d}B_s$ 是(B_t)鞅,再用推论 3.10,就得到 $\int_0^t \Phi_s \mathrm{d}B_s$ 是(B_t)鞅. 这就证明了结论(A).

(B)的证明. 容易想到随机变量 $\int_0^s \Phi_u \mathrm{d}B_u$ 是 $\boldsymbol{B}_s$ 可知的. 故而对于 $s<t$,利用已经证明了的结论(A),我们得到

$$\begin{aligned}E\left(\left[\int_0^s \Phi_u \mathrm{d}B_u \int_s^t \Phi_u \mathrm{d}B_u\right] \,\middle|\, \boldsymbol{B}_s\right)&=\int_0^s \Phi_u \mathrm{d}B_u E\left(\int_s^t \Phi_u \mathrm{d}B_u \,\middle|\, \boldsymbol{B}_s\right)\\&=\left(\int_0^s \Phi_u \mathrm{d}B_u\right)E\left(\int_s^t \Phi_u \mathrm{d}B_u \,\middle|\, \boldsymbol{B}_s\right)=0.\end{aligned}$$

另一方面,由引理 5.15 得到

$$\begin{aligned}E\left(\left(\int_0^t \Phi_u \mathrm{d}B_u\right)^2 \,\middle|\, \boldsymbol{B}_s\right)&=E\left(\left[\int_0^s \Phi_u \mathrm{d}B_u+\int_s^t \Phi_u \mathrm{d}B_u\right]^2 \,\middle|\, \boldsymbol{B}_s\right)\\&=\left(\int_0^s \Phi_u \mathrm{d}B_u\right)^2+E\left(\left(\int_s^t \Phi_u \mathrm{d}B_u\right)^2 \,\middle|\, \boldsymbol{B}_s\right)\\&=\left(\int_0^s \Phi_u \mathrm{d}B_u\right)^2+E\left(\int_s^t \Phi_u^2 \mathrm{d}u \,\middle|\, \boldsymbol{B}_s\right).\end{aligned}$$

于是

$$\begin{aligned}E(\eta_t\mid \boldsymbol{B}_s)&=E\left(\left(\int_0^t \Phi_u \mathrm{d}B_u\right)^2-\int_0^t \Phi_u^2 \mathrm{d}u \,\middle|\, \boldsymbol{B}_s\right)\\&=\left(\int_0^s \Phi_u \mathrm{d}B_u\right)^2+\int_s^t \Phi_u^2 \mathrm{d}u-\int_0^t \Phi_u^2 \mathrm{d}u=\eta_s.\end{aligned}$$

这就证明了(B).

至于结论(C)中$\{\zeta_t,t\geqslant 0\}$是(B_t)鞅的证明,我们将在下一段的例 5.8 中用 Ito 公式给出一个简洁的直观证明. 而另一个结论的证明,则需要用鞅的 Kolmogorov 不等式:

$$\begin{aligned}&P\left(\max_{t\leqslant 1}\left\{\int_0^t \Phi_s \mathrm{d}B_s-\frac{\alpha}{2}\int_0^t \Phi_s^2 \mathrm{d}s\right\}>\varepsilon\right)\\&=P\left(\max_{t\leqslant 1}\left\{\exp\left(\int_0^t \alpha\Phi_s \mathrm{d}B_s-\frac{1}{2}\int_0^t (\alpha\Phi_s)^2 \mathrm{d}s\right)\right\}>\mathrm{e}^{\alpha\varepsilon}\right)\end{aligned}$$

$$\leqslant \mathrm{e}^{-\alpha\varepsilon}E\left(\exp\left(\int_0^1 \alpha\Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^1(\alpha\Phi_s)^2\mathrm{d}s\right)\right)$$
$$\stackrel{\text{def}}{=\!=} \mathrm{e}^{-\alpha\varepsilon}E\zeta_1^{(\alpha)} = \mathrm{e}^{-\alpha\varepsilon}E\zeta_0^{(\alpha)} = \mathrm{e}^{-\alpha\varepsilon}.$$

注　用较为细致的鞅论技巧，还可以得到如下的深刻结论：

若 $E|\Phi_s|^2\equiv 1(s\geqslant 0)$，则 $\left\{\xi_t = \int_0^t \Phi_s \mathrm{d}B_s, t\geqslant 0\right\}$ 也是一个 Brown 运动.

在分析以 Brown 运动为随机干扰的随机系统的输出的性质的时候，我们还需要引入下面的概念.

(2) $(\boldsymbol{B}_t)$ 可知的随机过程按轨道的普通积分

令

$$\mathcal{L}^1 \stackrel{\text{def}}{=\!=} \left\{(B_t)\text{ 可知的随机过程 }\Phi_t\text{，且}\int_0^T E\mid\Phi_t(\omega)\mid \mathrm{d}t < +\infty\right\}.$$

在 $\mathcal{L}^1$ 中也可以定义如下的距离

$$d(\Psi,\Theta) \stackrel{\text{def}}{=\!=} \int_0^T E\mid \Psi_t - \Theta_t\mid \mathrm{d}t. \tag{5.20}$$

对 $\mathcal{L}^1$ 中的元 $\Psi \stackrel{\text{def}}{=\!=} \{\Psi_t: t\geqslant 0\}$，可以按样本轨道 ω 固定后定义普通积分：即对于固定的基本事件 ω 的积分 $\int_0^T \Psi_t(\omega)\mathrm{d}t$，可以证明它满足：

$$P\left(\omega: \int_0^T \Psi_t(\omega)\mathrm{d}t\text{ 存在}\right) = 1,$$
$$E\left(\int_0^T \Psi_t(\omega)\mathrm{d}t\right) = \int_0^T (E\Psi_t)\mathrm{d}t.$$

在 Ψ_t 都是连续的随机过程时，这些积分是普通的定积分，在一般的情形，也是定积分的推广，即 Lebesgue 积分.

5.3　Ito 公式——随机积分的换元公式与复合函数的随机微分公式

在普通函数的微积分中，复合函数的微分公式与积分的换元公式，是相互等价的两个最基本的公式. 而在随机微积分中，我们也有一个对应的公式，虽然它比微积分中的公式复杂，但是通过它仍然可以如在微积分中那样，将 Ito 积分用于许多问题中.

5.3.1　特殊类型的 Ito 过程

定义 5.6(特殊类型的 Ito 过程)　设

$$\xi_t = x + \int_0^t \Phi_s \mathrm{d}B_s + \int_0^t \Psi_s \mathrm{d}s, \tag{5.21}$$

其中随机过程 Φ_t,Ψ_t 分别是$\mathcal{L}^2$和$\mathcal{L}^1$中的元. 为了理解得更直接一些，我们还假定它们都是连续的随机过程. 即对任意固定的 ω，都是 t 的连续函数，这时 ξ_t 称为初值为 x，系数为 Φ_t,Ψ_t 的**特殊类型的 Ito 过程.**

我们常用如下的 **Ito(形式)微分**

$$d\xi_t = \Phi_t dB_t + \Psi_t dt. \tag{5.21$'$}$$

简单地表示 Ito 过程.

设 ξ_t 是 Ito 过程. 又二元实函数 $f(t,x)$ 对 x 二阶光滑且对 t 一阶光滑. 令 η_t 为复合得到的随机过程 $\eta_t = f(t,\xi_t)$，则下面的 Ito 公式表明 η_t 也是个 Ito 过程. 也就是说，Ito 过程对于这种光滑函数的复合运算是封闭的.

首先让我们仔细分析例 5.1 的结论：

$$\frac{1}{2}B_t^2 = \int_0^t B_s dB_s + \int_0^t \frac{1}{2}ds.$$

它说明了$\frac{1}{2}B_t^2$ 是一个 Ito 过程，且 $d\left(\frac{1}{2}B_t^2\right) = B_t dB_t + \frac{1}{2}dt$. 即对于 $f(x)=\frac{1}{2}x^2$ 而言，有

$$df(B_t) = d\left(\frac{1}{2}B_t^2\right) \neq B_t dB_t = f'(B_t)dB_t.$$

可见为了得到 $df(B_t)$，仅用类似于通常的微积分中 dB_t 的一次项 $f'(B_t)dB_t$ 展开是不够的，必须还补充以 $f(x)=\frac{1}{2}x^2$ 的 Taylor 展开中的第二项$\frac{1}{2}f''(B_t)(dB_t)^2 = \frac{1}{2}(dB_t)^2$，它提供了 $d\left(\frac{1}{2}B_t^2\right)$中的第二项$\frac{1}{2}dt$.

引理 5.17

$$(dB_t)^2 = dt. \tag{5.22}$$

在国际上一些金融衍生证券的教材中，这个公式常通俗地写为

$$dB_t = \sqrt{dt}. \tag{5.22$'$}$$

证明 由例 5.1

$$d\left(\frac{1}{2}B_t^2\right) = B_s dB_s + \frac{1}{2}dt,$$

另一方面，由 Taylor 展开

$$d\left(\frac{1}{2}B_t^2\right) = B_s dB_s + \frac{1}{2}(dB_t)^2.$$

将两者进行比较便得到结论.

这个引理说明，如果在 $dt\to 0$ 时，将 dt 视为一阶无穷小，那么，dB_t 就说是半阶无穷小 $\sqrt{dt}$. 由此启示我们，对 Ito 过程$\{\xi_t: t\geqslant 0\}$的复合函数 $f(\xi_t)$作微分(即随机微分)的时候，必须将 Taylor 展开到二阶项，即用

$$\mathrm{d}f(\xi_t) = f'(\xi_t)\mathrm{d}\xi_t + \frac{1}{2}f''(\xi_t)(\mathrm{d}\xi_t)^2,$$

这样才能包含了半阶无穷小所作出的贡献.

将引理 5.17 的结论一般化，就是 Ito 公式. 它说明对于很好的函数 $f(x)$，例如，f'，f'' 都有界的函数(从而保证 $f'(\xi_t)$ 和 $f''(\xi_t)$ 分别属于 $\mathcal{L}^1$ 和 $\mathcal{L}^2$)，就可容易地用 Taylor 展开式证明 Ito 过程 ξ_t 的复合函数 $f(\xi_t)$ 仍是 Ito 过程.

然而，要求 f'，f'' 有界过于苛刻，即使以上的函数 $\frac{1}{2}x^2$ 和指数函数等初等函数都不能满足，而不能对更多的 f 建立 Ito 公式的其原因在于 Ito 积分中对被积函数属于 $\mathcal{L}^2$ 的要求过于局限. 我们需要对于 f'，f'' 并不有界的函数，建立复合函数 $f(\xi_t)$ 公式，首先就需要推广 Ito 积分的被积函数类. 我们要求使展开式 $\mathrm{d}f(\xi_t)=f'(\xi_t)\mathrm{d}\xi_t+\frac{1}{2}f''(\xi_t)(\mathrm{d}\xi_t)^2$ 中涉及的积分的可积性质仅由函数 $f(x)$ 的可微性质(这类条件容易验证)所保证，而不再加上诸如 $f'(\xi_t)$ 和 $f''(\xi_t)$ 属于 $\mathcal{L}^2$ 或 $\mathcal{L}^1$ 等假定(因为这类条件很难验证).

为了达到这个目的，我们先需要将 Ito 积分作必要的推广.

5.3.2　Ito 积分中被积随机过程类的推广与一般的 Ito 过程

记

$$\mathcal{L}^{2,\mathrm{loc}} \stackrel{\mathrm{def}}{=\!=} \left\{(B_t)\text{ 可知的随机过程 }\Phi_t\text{，且 }P\left(\omega\colon \int_0^T \Phi_t(\omega)^2\mathrm{d}t < +\infty\right) = 1\right\}.$$

可以证明如下事实：

(1) 对于任意 $\Phi\in\mathcal{L}^{2,\mathrm{loc}}$，存在 (B_t) 停时序列 $\tau_1\leqslant\cdots\leqslant\tau_n\uparrow+\infty$，使得对于固定的 n，随机过程 $\Phi_t^{(n)} \stackrel{\mathrm{def}}{=\!=} \Phi_{t\wedge\tau_n}\in\mathcal{L}^2$.

(2) $\int_0^{t\wedge\tau_n}\Phi_t^{(n+1)}\mathrm{d}B_t = \int_0^{t\wedge\tau_n}\Phi_t^{(n)}\mathrm{d}B_t$ 从而概率为 1 地存在极限，于是，可以定义

$$\int_0^t \Phi_s\mathrm{d}B_s \stackrel{\mathrm{def}}{=\!=} \text{概率为 1 地 }\lim_{n\to+\infty}\int_0^t\Phi_s^{(n)}\mathrm{d}B_s$$

(因此也依概率收敛). 以上定义说明了 $\mathcal{L}^{2,\mathrm{loc}}$ 过程的 Ito 积分正是 $\mathcal{L}^2$ 过程的 Ito 积分沿轨道的延伸.

(3) $\xi_t \stackrel{\mathrm{def}}{=\!=} \int_0^t\Phi_s\mathrm{d}B_s$ 是 (B_t) 局部鞅，即对于固定的 n，$\zeta_t^{(n)} \stackrel{\mathrm{def}}{=\!=} \xi_{t\wedge\tau_n}$ 是 (B_t) 鞅.

再记

$$\mathcal{L}^{1,\mathrm{loc}} \stackrel{\mathrm{def}}{=\!=} \left\{(B_t)\text{ 可知的随机过程 }\Psi_t\text{，且 }P\left(\omega\colon \int_0^T |\Psi_t(\omega)|\,\mathrm{d}t < +\infty\right) = 1\right\}.$$

那么类似地，对于任意 $\Psi\in\mathcal{L}^{1,\mathrm{loc}}$，积分 $\int_0^t\Psi_s(\omega)\mathrm{d}s$ 就有意义.

定义 5.6′(一般的 Ito 过程) 设

$$\xi_t = x + \int_0^t \Phi_s \mathrm{d}B_s + \int_0^t \Psi_s \mathrm{d}s,$$

其中随机过程 Φ_t, Ψ_t 分别是$\mathcal{L}^{2,\mathrm{loc}}$和$\mathcal{L}^{1,\mathrm{loc}}$中的元,则 ξ_t 称为初值为 x,系数为 Φ_t, Ψ_t 的**一般的 Ito 过程**,简称 Ito 过程.

形式地,对 Ito 过程$\{\xi_t: t\geqslant 0\}$的 Ito 积分自然地可以理解为

$$\int_0^T \Theta_t \mathrm{d}\xi_t \xlongequal{\mathrm{def}} \int_0^T \Theta_t \Phi_t \mathrm{d}B_t + \int_0^T \Theta_t \Psi_t \mathrm{d}t.$$

定理 5.18(Ito 公式,或随机微分公式) 设 ξ_t 为 Ito 过程,即

$$\mathrm{d}\xi_t = \Phi_t \mathrm{d}B_t + \Psi_t \mathrm{d}t.$$

又实值函数 $f(t,x)$对 x 二次连续可微,且对 t 连续可微. 那么,$\eta_t = f(t,\xi_t)$也是 Ito 过程,而且

$$\begin{aligned}
\mathrm{d}\eta_t &= \mathrm{d}f(t,\xi_t) + \frac{1}{2}\mathrm{d}^2 f(t,\xi_t) \\
&= f'_t(t,\xi_t)\mathrm{d}t + f'_x(t,\xi_t)\mathrm{d}\xi_t + \frac{1}{2} f''_{xx}(t,\xi_t)(\mathrm{d}\xi_t)^2 \\
&= \left(f'_t + f'_x \Psi_t + \frac{1}{2}\Phi_t^2 f''_{xx}\right)\mathrm{d}t + f'_x \Phi_t \mathrm{d}B_t,
\end{aligned} \tag{5.23}$$

其中

$$(\mathrm{d}\xi_t)^2 = (\Phi_t \mathrm{d}B_t + \Psi_t \mathrm{d}t)^2 \xlongequal{\mathrm{def}} \Phi_t^2 \mathrm{d}t. \tag{5.24}$$

证明 理解和证明 Ito 公式的核心点是引理 5.17. Ito 公式实际上就是 $f(t,\xi_t)$的 Taylor 展开式关于 $\mathrm{d}t$ 的一阶近似. Taylor 展开实际上给出了 Ito 公式的直观推导. 其严格推导则需要较为精致的随机分析工具,本书不再列入.

推论 5.19 设$\{\xi_t: t\geqslant 0\}$为 Ito 过程: $\mathrm{d}\xi_t = \Phi_t \mathrm{d}B_t + \Psi_t \mathrm{d}t$. $f(x)$二阶连续可微. 那么,$\{f(\xi_t): t\geqslant 0\}$也是 Ito 过程,而且有

$$\begin{aligned}
\mathrm{d}f(\xi_t) &= \mathrm{d}f(\xi_t) + \frac{1}{2}\mathrm{d}^2 f(\xi_t) = f'(\xi_t)\mathrm{d}\xi_t + \frac{1}{2}f''(\xi_t)(\mathrm{d}\xi_t)^2 \\
&= \left(f'\Psi_t + \frac{1}{2}\Phi_t^2 f''\right)\mathrm{d}t + f'_x \Phi_t \mathrm{d}B_t.
\end{aligned} \tag{5.23$'$}$$

5.3.3 多维 Brown 运动积分的 Ito 过程

定义 5.7 m 个相互独立的 Brown 运动组成的向量 $\begin{pmatrix} B_t^{(1)} \\ B_t^{(2)} \\ \vdots \\ B_t^{(m)} \end{pmatrix}$,称为 m 维 Brown 运动,或$\mathbb{R}^m$ 中的 Brown 运动,记成 $\boldsymbol{B}_t$,即

$$\boldsymbol{B}_t \xlongequal{\text{def}} \begin{pmatrix} B_t^{(1)} \\ B_t^{(2)} \\ \vdots \\ B_t^{(m)} \end{pmatrix}.$$

注意，前面的叙述中，曾将 $\boldsymbol{B}_t$ 理解为 Brown 运动在时间 t 前的信息 $\{B_s: s\leqslant t\}$，而这里又将它记为向量过程. 根据不同的定语所提供的信息，并不会引起误解.

定义 5.8（多维 Brown 运动的 Ito 过程） 设

$$\xi_t = x + \int_0^t \boldsymbol{\Phi}_s^T \mathrm{d}\boldsymbol{B}_s + \int_0^t \Psi_s \mathrm{d}s,$$

其中

$$\boldsymbol{\Phi}_s = \begin{pmatrix} \Phi_t^{(1)} \\ \Phi_t^{(2)} \\ \vdots \\ \Phi_t^{(m)} \end{pmatrix},$$

而随机过程 $\Phi_t^{(k)}(k\leqslant m)$，$\Psi_t$ 分别是 $\mathcal{L}^{2,\text{loc}}$ 和 $\mathcal{L}^{1,\text{loc}}$ 中的元. 那么，ξ_t 也称为初值为 x 的**多维 Brown 运动的 Ito 过程**.

推论 5.19′ 设实值函数 $f(t,x)$ 对 x 二次连续可微且对 t 连续可微. 那么，Ito 公式对于多维 Brown 运动的 Ito 过程

$$\mathrm{d}\xi_t = \sum_{i=1}^m \Phi_t^{(i)} \mathrm{d}B_t^{(i)} + \Psi_t \mathrm{d}t$$

仍然适用，即有

$$\mathrm{d}f(t,\xi_t) = f'_t \mathrm{d}t + f'_x \mathrm{d}\xi_t + \frac{1}{2} f''_{xx} (\mathrm{d}\xi_t)^2, \tag{5.23″}$$

其中

$$(\mathrm{d}\xi_t)^2 = \sum_{i=1}^m (\Phi_t^{(i)})^2 \mathrm{d}t. \tag{5.24′}$$

命题 5.20 设 $\{\xi_t: t\geqslant 0\}$ 为 Ito 过程：$\mathrm{d}\xi_t = \Phi_t \mathrm{d}B_t + \Psi_t \mathrm{d}t$. $f(x)$ 二次连续可微. 那么 Stratonovich 积分于 Ito 积分之间有如下联系：

$$\int_0^t f(\xi_s) \circ \mathrm{d}B_s = \int_0^t f(\xi_s) \mathrm{d}B_s + \frac{1}{2} \int_0^t f'(\xi_s) \Phi_s \mathrm{d}s, \tag{5.25}$$

即

$$f(\xi_t) \circ \mathrm{d}B_t = f(\xi_t) \mathrm{d}B_t + \frac{1}{2} \mathrm{d}f(\xi_t) \mathrm{d}B_t. \tag{5.25′}$$

其中

$$\mathrm{d}f(\xi_t) \mathrm{d}B_t = (\Phi_t \mathrm{d}B_t)(\mathrm{d}B_t) = \Phi_t \mathrm{d}t. \tag{5.26}$$

证明 类似于引理 5.13 的证明,我们可以得到

$$(p)\lim_{n\to+\infty}\sum f'(\xi_{t_k^{(n)}})\Phi_{t_i^{(n)}}(B_{t_{k+1}^{(n)}\wedge t}-B_{t_k^{(n)}\wedge t})^2=\int_0^t f'(\xi_s)\Phi_s\mathrm{d}s.$$

于是

$$\begin{aligned}\int_0^t f(\xi_s)\circ\mathrm{d}B_s&=(p)\lim_{n\to+\infty}\sum f\left(\frac{\xi_{t_{k+1}^{(n)}}+\xi_{t_k^{(n)}}}{2}\right)(B_{t_{k+1}^{(n)}\wedge t}-B_{t_k^{(n)}\wedge t})\\&=(p)\lim_{n\to+\infty}\sum f(\xi_{t_k^{(n)}})(B_{t_{k+1}^{(n)}\wedge t}-B_{t_k^{(n)}\wedge t})\\&\quad+(p)\lim_{n\to+\infty}\sum\left[f\left(\frac{\xi_{t_{k+1}^{(n)}}+\xi_{t_k^{(n)}}}{2}\right)-f(\xi_{t_k^{(n)}})\right](B_{t_{k+1}^{(n)}\wedge t}-B_{t_k^{(n)}\wedge t})\\&=\int_0^t f(\xi_s)\mathrm{d}B_s+(p)\lim_{n\to+\infty}\frac{1}{2}\sum f'(\theta_k^{(n)})\Phi_{t_i^{(n)}}(B_{t_{k+1}^{(n)}\wedge t}-B_{t_k^{(n)}\wedge t})^2=\text{右},\end{aligned}$$

其中 $\theta_k^{(n)}$ 介于 $B_{t_k^{(n)}}$ 和 $B_{t_{k+1}^{(n)}}$ 之间.

命题 5.21 设$\{\xi_t: t\geqslant 0\}$是 Stratonovich 型的 Ito 过程,即

$$\xi_t=x+\int_0^t\Phi_s\circ\mathrm{d}B_s+\int_0^t\Psi_s\mathrm{d}s\quad(\text{写成 }\mathrm{d}\xi_t=\Phi_t\circ\mathrm{d}B_t+\Psi_t\mathrm{d}t),\tag{5.27}$$

那么,它有如同普通的微积分那样的复合函数的微分的链法则;即对于三次连续可微的函数 $f(x)$有

$$\mathrm{d}f(\xi_t)=f'(\xi_t)\circ\mathrm{d}\xi_t.\tag{5.28}$$

而对 x 三次连续可微,对 t 连续可微 $f(t,x)$,有

$$\mathrm{d}f(t,\xi_t)=f'_t\mathrm{d}t+f'_x\circ\mathrm{d}\xi_t.\tag{5.28$'$}$$

证明 由推论 5.19 和命题 5.20,有

$$\begin{aligned}\mathrm{d}f(\xi_t)&=f'(\xi_t)\mathrm{d}\xi_t+\frac{1}{2}f''(\xi_t)(\mathrm{d}\xi_t)^2\\&=f'(\xi_t)\mathrm{d}\xi_t+\frac{1}{2}[\mathrm{d}f'(\xi_t)]\mathrm{d}\xi_t=f'(\xi_t)\circ\mathrm{d}\xi_t.\end{aligned}$$

而对于函数 $f(t,x)$,则有

$$\begin{aligned}\mathrm{d}f(t,\xi_t)&=f'_t(t,\xi_t)\mathrm{d}t+f'_x(t,\xi_t)\mathrm{d}\xi_t+\frac{1}{2}f''_{xx}(\mathrm{d}\xi_t)^2\\&=f'_t\mathrm{d}t+f'_x\mathrm{d}\xi_t+\frac{1}{2}\mathrm{d}f'_x\mathrm{d}\xi_t=f'_t+f'_x\circ\mathrm{d}\xi_t.\end{aligned}$$

注 命题 5.21 进一步说明了 Stratonovich 积分的复合运算规律与普通微积分一致.

下面讨论 Stratonovich 微分的运算规律.

定理 5.22 对于如下定义的 Ito 过程(ξ_t),(η_t),(ζ_t):

$$\mathrm{d}\xi_t=\Phi_t^{(\xi)}\mathrm{d}B_t+\Psi_t^{(\xi)}\mathrm{d}t,\quad\mathrm{d}\eta_t=\Phi_t^{(\eta)}\mathrm{d}B_t+\Psi_t^{(\eta)}\mathrm{d}t,\quad\mathrm{d}\zeta_t=\Phi_t^{(\zeta)}\mathrm{d}B_t+\Psi_t^{(\zeta)}\mathrm{d}t,$$

我们有

(1) $d\xi_t dt=0$,

(2) $d\xi_t d\eta_t=\Phi_t^{(\xi)}\Phi_t^{(\eta)}dt$,

(3) $d\xi_t d\eta_t d\zeta_t=0$,

(4) $\xi_t\circ(d\eta_t d\zeta_t)=(\xi_t\circ d\eta_t)d\zeta_t=\xi_t(d\eta_t d\zeta_t)$,

(5) $\xi_t\circ(\eta_t\circ d\zeta_t)=(\xi_t\eta_t)\circ d\zeta_t$.

证明　(1) 直接得自定义.

(2) 的证明只需利用

$$d\xi_t d\eta_t=\frac{(d[\xi_t+\eta_t])^2-(d\xi_t)^2-(d\eta_t)^2}{2}.$$

(3) 得自

$$d\xi_t d\eta_t d\zeta_t=d\xi_t(d\eta_t d\zeta_t)=d\xi_t\left(\frac{1}{2}\Phi_t^{(\eta)}\Phi_t^{(\zeta)}dt\right)=0.$$

(4) 的证明：

$$\xi_t\circ(d\eta_t d\zeta_t)=\xi_t(d\eta_t d\zeta_t)+\frac{1}{2}d\xi_t(d\eta_t d\zeta_t)=\xi_t(d\eta_t d\zeta_t),$$

$$(\xi_t\circ d\eta_t)d\zeta_t=\left(\xi_t d\eta_t+\frac{1}{2}d\xi_t d\eta_t\right)d\zeta_t=\xi_t(d\eta_t d\zeta_t).$$

(5) 的证明：

$$\begin{aligned}\xi_t\circ(\eta_t\circ d\zeta_t)&=\xi_t\circ\left(\eta_t d\zeta_t+\frac{1}{2}d\eta_t d\zeta_t\right)\\&=\xi_t\left(\eta_t d\zeta_t+\frac{1}{2}d\eta_t d\zeta_t\right)+\frac{1}{2}d\xi_t\left(\eta_t d\zeta_t+\frac{1}{2}d\eta_t d\zeta_t\right)\\&=\xi_t\left(\eta_t d\zeta_t+\frac{1}{2}d\eta_t d\zeta_t\right)+\frac{1}{2}(d\xi_t)\eta_t d\zeta_t\\&=(\xi_t\eta_t)\left[d\zeta_t+\frac{1}{2}d(\xi_t\eta_t)d\zeta_t\right]=(\xi_t\eta_t)\circ d\zeta_t.\end{aligned}$$

定理 5.23(同一个 Brown 运动的多个 Ito 过程的 Ito 公式)　设 $\boldsymbol{\xi}_t=(\xi_t^{(1)},\xi_t^{(2)},\cdots,\xi_t^{(d)})$，$\xi_t^{(i)}(i\leqslant d)$都是同一个 Brown 运动的 Ito 过程：

$$d\xi_t^{(i)}=\Phi_t^{(i)}dB_t+\Psi_t^{(i)}dt.$$

$f(t,\boldsymbol{x})$对 $\boldsymbol{x}$ 二次连续可微，且对 t 连续可微. 那么，$\eta_t=f(t,\boldsymbol{\xi}_t)$也是 Ito 过程，而且有

$$\begin{aligned}d\eta_t&=df(t,\xi_t)+\frac{1}{2}d^2f(t,\xi_t)\\&=f'_t(t,\xi_t)dt+\sum_{j=1}^{d}f'_{x_i}(t,\boldsymbol{\xi}_t)d\xi_t^{(i)}+\frac{1}{2}\sum_{i,j=1}^{d}f''_{x_ix_j}(t,\boldsymbol{\xi}_t)(d\xi_t^{(i)}d\xi_t^{(j)})\\&=\left[f'_t+\sum_{i=1}^{d}f'_{x_i}\Psi_t^{(i)}+\frac{1}{2}\sum_{i,j=1}^{d}\Phi_t^{(i)}\Phi_t^{(j)}f''_{x_ix_j}\right]dt+\left[\sum_{i=1}^{d}f'_{x_i}\Phi_t^{(i)}\right]dB_t,\end{aligned}$$

其中

$$\mathrm{d}\xi_t^{(i)}\mathrm{d}\xi_t^{(j)} = (\Phi_t^{(i)}\mathrm{d}B_t + \Psi_t^{(i)}\mathrm{d}t)(\Phi_t^{(j)}\mathrm{d}B_t + \Psi_t^{(j)}\mathrm{d}t) = \Phi_t^{(i)}\Phi_t^{(j)}\mathrm{d}t.$$

推论 5.24(乘积公式) 设 ξ_t,η_t 都是同一个 Brown 运动的 Ito 过程,则

$$\mathrm{d}(\xi_t\eta_t) = \xi_t\mathrm{d}\eta_t + \eta_t\mathrm{d}\xi_t + (\mathrm{d}\xi_t)(\mathrm{d}\eta_t).$$

证明 这是 $f(x_1x_2)=x_1x_2$ 的情形.

命题 5.25 设 ξ_t 为 Ito 过程,$g(x)$为有界连续函数.又若 $f(t,x)$对 x 二次连续可微,且对 t 连续可微,$h(x)$连续可微.那么

$$\zeta_t \overset{\text{def}}{=\!=} h\left(\int_0^t g(\xi_s)\mathrm{d}s\right)f(t,\xi_t)$$

也是 Ito 过程,而且有

$$\mathrm{d}\zeta_t = f(t,\xi_t)h'\left(\int_0^t g(\xi_s)\mathrm{d}s\right)g(\xi_t)\mathrm{d}t + h\left(\int_0^t g(\xi_s)\mathrm{d}s\right)\mathrm{d}f(t,\xi_t).$$

证明 显见 $\eta_t \overset{\text{def}}{=\!=} \int_0^t g(\xi_s)\mathrm{d}s$ 也是 Ito 过程.对 $h(\eta_t)f(t,\xi_t)$用乘积 Ito 公式(注意 $\mathrm{d}\xi_t\mathrm{d}\eta_t=0$)直接得到

$$\begin{aligned}\mathrm{d}\zeta_t &= h'(\eta_t)f(t,\xi_t)\mathrm{d}\eta_t + h(\eta_t)\mathrm{d}f(t,\xi_t) + \mathrm{d}\xi_t\mathrm{d}\eta_t \\ &= h'(\eta_t)f(t,\xi_t)g(\xi_t)\mathrm{d}t + h(\eta_t)\mathrm{d}f(t,\xi_t).\end{aligned}$$

例 5.5

$$\mathrm{d}\left(\mathrm{e}^{-\int_0^t c(\xi_s)\mathrm{d}s}f(t,\xi_t)\right) = \mathrm{e}^{-\int_0^t c(\xi_s)\mathrm{d}s}[\mathrm{d}f(t,\xi_t) - c(\xi_t)f(t,\xi_t)\mathrm{d}t].$$

特别地,当 $f(t,x)=x$ 时有

$$\mathrm{d}\left(\mathrm{e}^{-\int_0^t c(\xi_s)\mathrm{d}s}\xi_t\right) = \mathrm{e}^{-\int_0^t c(\xi_s)\mathrm{d}s}[\mathrm{d}\xi_t - c(\xi_t)\xi_t\mathrm{d}t].$$

例 5.6

$$\int_0^t B_s^n\mathrm{d}B_s = \frac{1}{n+1}B_t^{n+1} - \frac{n}{2}\int_0^t B_s^{n-1}\mathrm{d}s.$$

特别地

$$\int_0^t B_s^2\mathrm{d}B_s = \frac{1}{3}B_t^3 - \int_0^t B_s\mathrm{d}s.$$

证明 对 $f(B_t)\overset{\text{def}}{=\!=}B_t^n(n\geqslant 2)$,用 Ito 公式,得到

$$\mathrm{d}B_t^n = f'(B_t)\mathrm{d}B_t + \frac{1}{2}f''(B_t)(\mathrm{d}B_t)^2 = nB_t^{n-1}\mathrm{d}B_t + \frac{n(n-1)}{2}B_t^{n-2}\mathrm{d}t,$$

即

$$\int_0^t B_s^{n-1}\mathrm{d}B_s = \frac{1}{n}B_t^n - \frac{(n-1)}{2}\int_0^t B_s^{n-2}\mathrm{d}s.$$

例 5.7(Ito 积分的指数鞅) 若(B_t)可知随机过程 Φ_t 有界,即$|\Phi_t|<$某个常数,则

$$\zeta_t \overset{\text{def}}{=\!=} \mathrm{e}^{\int_0^t \Phi_s\mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2\mathrm{d}s}$$

是(B_t)鞅.

直观证明　对

$$\zeta_t \stackrel{\text{def}}{=\!=} f\left(\int_0^t \Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s\right) \stackrel{\text{def}}{=\!=} \mathrm{e}^{\int_0^t \Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s}$$

用 Ito 公式得到

$$\mathrm{d}\zeta_t = \mathrm{e}^{\int_0^t \Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s}\left[\left(\Phi_t \mathrm{d}B_t - \frac{1}{2}\Phi_t^2 \mathrm{d}t\right) + \frac{1}{2}\Phi_t^2 (\mathrm{d}B_t)^2\right] = \mathrm{e}^{\int_0^t \Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s}\Phi_t \mathrm{d}B_t.$$

可见 ζ_t 是(B_t)鞅. 因为我们暂时不能证明被积随机过程 $\mathrm{e}^{\int_0^t \Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s}\Phi_t$ 属于$\mathcal{L}^2$，所以这里只是直观的证明. 利用在停时上停止的技术和极限过渡，就可以严格地证明此断言.

Ito 公式是随机微积分的核心，它的成立的范围可以更广，例如复合函数 f 不必二阶连续可微，而只是一个凸函数的情形，此时在 Ito 公式中将会出现一个附加的单调的随机过程，称为**局部时**，粗略地讲，它体现了随机过程在二次导数的不连续点上的时间，因为这将涉及较深的理论，我们不再介绍. 本书只就最通常的情形，弄清 Ito 公式的实质，并通过它理解与掌握随机微积分的要领.

例 5.8(Ito 过程的幂函数复合)　设 $\eta_t = \xi_t^\alpha$，其中 ξ_t 是如下的正值 Ito 过程

$$\mathrm{d}\xi_t = \Phi_s \mathrm{d}B_s + \Psi_s \mathrm{d}s.$$

那么，

$$\mathrm{d}\eta_t = \alpha\xi_t^{\alpha-1}\mathrm{d}\xi_t + \frac{\alpha(\alpha-1)}{2}\xi_t^{\alpha-2}\Phi_t^2 \mathrm{d}t.$$

特别地，若 $\alpha = \dfrac{1}{2}$，则

$$\mathrm{d}\sqrt{\xi_t} = \frac{1}{2\sqrt{\xi_t}}\mathrm{d}\xi_t - \frac{1}{8\xi_t\sqrt{\xi_t}}\Phi_t^2 \mathrm{d}t.$$

例 5.9　对于连续可微函数 $g(t,x)$，直接验证可得

$$\mathrm{d}\left(\int_0^t g(t,s)\mathrm{d}B_s\right) = g(t,t)\mathrm{d}B_t + \left(\int_0^t \frac{\partial g}{\partial t}\mathrm{d}B_s\right)\mathrm{d}t.$$

* **定理 5.26**(多维 Brown 运动的多维 Ito 公式)　设$\boldsymbol{\xi}_t = (\xi_t^{(1)}, \xi_t^{(2)}, \cdots, \xi_t^{(d)})$，$\xi_t^{(i)}$ $(i \leqslant d)$都是多维 Brown 运动$\{\mathrm{d}\boldsymbol{B}_t: t \geqslant 0\}$的 Ito 过程：

$$\mathrm{d}\xi_t^{(i)} = \boldsymbol{\Phi}_t^{(i)\mathrm{T}}\mathrm{d}\boldsymbol{B}_t + \Psi_t^{(i)}\mathrm{d}t.$$

$f(t,\boldsymbol{x})$为对 $\boldsymbol{x}$ 二阶连续可微，且对 t 连续可微. 那么，$\eta_t = f(t,\boldsymbol{\xi}_t)$也是 Ito 过程，而且仍然有

$$\begin{aligned}\mathrm{d}\eta_t &= \mathrm{d}f(t,\xi_t) + \frac{1}{2}\mathrm{d}^2 f(t,\xi_t)\\ &= f_t'(t,\boldsymbol{\xi}_t)\mathrm{d}t + \sum_{j=1}^d f_{x_i}'(t,\boldsymbol{\xi}_t)\mathrm{d}\xi_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^d f_{x_i x_j}''(t,\boldsymbol{\xi}_t)(\mathrm{d}\xi_t^{(i)}\mathrm{d}\xi_t^{(j)}).\end{aligned}$$

其中

$$\mathrm{d}\xi_t^{(i)}\mathrm{d}\xi_t^{(j)} = \boldsymbol{\Phi}_t^{(i)\mathrm{T}}\boldsymbol{\Phi}_t^{(j)}\mathrm{d}t.$$

推论 5.24′(乘积公式) 设 ξ_t,η_t 都是多维 Brown 运动$\{\boldsymbol{B}_t: t\geqslant 0\}$的 Ito 过程,则

$$\mathrm{d}(\xi_t\eta_t) = \xi_t\mathrm{d}\eta_t + \eta_t\mathrm{d}\xi_t + (\mathrm{d}\xi_t)(\mathrm{d}\eta_t).$$

推论 5.21′(多个 Brown 运动情形 Stratonovich 积分型的 Ito 过程的链法则) 假设 m 维 Brown 运动

$$\boldsymbol{B}_t = \begin{pmatrix} B_t^{(1)} \\ B_t^{(2)} \\ \vdots \\ B_t^{(m)} \end{pmatrix},$$

$$\xi_t = x + \int_0^t \boldsymbol{\Phi}_s^{\mathrm{T}} \circ \mathrm{d}\boldsymbol{B}_s + \int_0^t \Psi_s \mathrm{d}s,$$

其中

$$\boldsymbol{\Phi}_s = \begin{pmatrix} \Phi_t^{(1)} \\ \Phi_t^{(2)} \\ \vdots \\ \Phi_t^{(m)} \end{pmatrix},$$

$\Phi_t^{(k)}\,(k\leqslant m)$,$\Psi_t$ 分别属于$\mathcal{L}^{2,\mathrm{loc}}$和$\mathcal{L}^{1,\mathrm{loc}}$的($\xi_t$ 是初值为 x 的多维 Brown 运动的 Stratonovich 积分型的 Ito 过程). 那么,对于光滑函数 $f(t,x)$有

$$\mathrm{d}f(t,\xi_t) = f'_t\mathrm{d}t + f'_x \circ \mathrm{d}\xi_t.$$

这个公式也可以推广到 d 维过程$\boldsymbol{\xi}_t$ 的情形. 特别地,如果

$$\boldsymbol{\xi}_t = \boldsymbol{x} + \int_0^t \boldsymbol{\Sigma}(t,\boldsymbol{\xi}_s)^{\mathrm{T}} \circ \mathrm{d}\boldsymbol{B}_s + \int_0^t \boldsymbol{b}(t,\boldsymbol{\xi}_s)^{\mathrm{T}}\mathrm{d}s.$$

其中$\boldsymbol{\Sigma}(t,\bar{\boldsymbol{x}})=(\sigma_{i,j}(t,\boldsymbol{x}))_{i\leqslant d,j\leqslant m}$,$\boldsymbol{b}(t,\bar{\boldsymbol{x}})=(b_i(t,\boldsymbol{x}))_{i\leqslant d}$. 这时有

$$\mathrm{d}f(t,\boldsymbol{\xi}_t) = f'_t\mathrm{d}t + \sum_{i=1}^d f'_{x_i} \circ \mathrm{d}\xi_t^{(i)}.$$

*5.3.4 用鞅的语言来表达 Ito 公式——Stroock-Varadhan 表示

在 20 世纪 60 年代末至 70 年代初,Stroock-Varadhan 首先意识到,在描述宏观的扩散现象时,涉及的是 Brown 运动的轨道函数,或 Ito 积分的轨道函数(一般统称为 **Brown 泛函**)的统计性质,而求统计平均(数学期望)时,其中的 Brown 运动只是一个代表微观的随机涨落的驱动力,不会真正在宏观的平均性质中出现. 所以在模型的微观表达式中,重要的只是需要有一个 Brown 运动,至于是哪一个 Brown 运动,则是无关紧要的,因为,不同的 Brown 运动的相同形式的泛函,其平均是一样的. 也就是说,这里的 Brown 运动,只

是一个**参变的随机过程**. 与解析几何中从曲线的参数方程消去参数后得到曲线的普通方程的想法一样，在求宏观的平均时，可以消去参变的 Brown 运动.

于是"$\{\xi_t:\ t\geqslant 0\}$为 Ito 过程 $d\xi_t=\Phi_t dB_t+\Psi_t dt$"可以叙述为"$\xi_t-\int_0^t \Psi_s ds$ 是一个鞅"(数学上更严格地说，是局部鞅，我们并不细究). 而 Ito 公式则可以叙述为对于可微性质较好的函数 $f(t,x)$,

$$f(t,\xi_t)-\int_0^t\left(f'_t+f'_x\Psi_s+\frac{1}{2}\Phi_s^2 f''_{xx}\right)_{(t,\xi_t)}ds$$

是一个鞅.

反之，如果对于足够多的性质好的函数，如下定义的随机过程

$$M_t^{(f)}\stackrel{\text{def}}{=\!=}f(t,\xi_t)-\int_0^t\left(f'_t+f'_x\Psi_s+\frac{1}{2}\Phi_s^2 f''_{xx}\right)_{(t,\xi_t)}ds$$

都是鞅，则如下的**鞅表示定理**成立，即存在一个 Brown 运动 B_t，使

$$d\xi_t=\Phi_t dB_t+\Psi_t dt.$$

也就是说，$\{\xi_t:\ t\geqslant 0\}$是 Ito 过程.

习题 5

1. 对于数值函数 $f(t)$, $g(t)$，求 Ito 过程 $\xi_t=\int_0^t f(u)dB_u+\int_0^t g(u)du$ 的 n 阶中心矩 $E(\xi_t-E\xi_t)^n$，并求在 $\xi_s=x$ 条件下的条件分布密度.

2. 用 Ito 公式求 $e^{B(t)}$ 的期望.

3. 用 Ito 公式求 $E\left(\int_0^t \Phi_s dB_s\right)^{2n}$，并证明存在只依赖于 n 的常数 C_n 使

$$E\left(\int_0^t \Phi_s dB_s\right)^{2n}\leqslant C_n\int_0^t E\Phi_s^{2n}ds.$$

4. 设 $f(t)$ 为平方可积的实值函数，求证 $X_t=\int_0^t f(s)dB_s$ 为 Gauss 过程，并计算 $\mathrm{cov}(X_s,X_t)$. 进一步证明 (X_t,Y_t) 是二维 Gauss 过程，其中 $Y_t=\int_0^t g(t)X_t dt$. 又问 $\mathrm{cov}(X_s,Y_t)$, $\mathrm{cov}(Y_s,Y_t)$ 各是什么?

5. 设 B_t 为 Brown 运动，$f(x)$ 为二次连续可微的增函数. 证明 $\xi_t=e^{af(B_t)-\frac{1}{2}\int_0^t[a^2 f'(B_s)^2+af''(B_s)]ds}$ 是 (B_t) 鞅.

6. 假设 ξ_t 是 Ito 过程，又设 $f(t,x)$ 为对 x 二阶光滑(二阶导数连续)且对 t 一阶光滑(一阶导数连续)的实函数，$\eta_t=e^{\int_0^t \xi_s ds}f(t,\xi_t)$. 求 $d\eta_t$.

7. 由定义求 $\int_0^t s\mathrm{d}B_s$，$\int_0^t B_s^2\mathrm{d}B_s$.

8. 求由原点出发的二维 Brown 运动在时刻 t 落在圆盘 $x^2+y^2\leqslant\rho^2$ 中的概率.

9. 判别下列是否为鞅(需要证明)：$t^2B_t-2\int_0^t sB_s\mathrm{d}s$，$B_t^3-3tB_t$.

10. 设 $(B_t^{(1)}, B_t^{(2)})$ 是二维 Brown 运动. 判断并证明下列是否为鞅：$B_t^{(1)}B_t^{(2)}$，$\ln([B_t^{(1)}]^2+[B_t^{(2)}]^2)$.

11. 将下列 Ito 方程化为 Stratonovich 方程：$\mathrm{d}X_t=\beta X_t\mathrm{d}t+\alpha X_t\mathrm{d}B_t$，$\mathrm{d}X_t=2\mathrm{e}^{-X_t}\mathrm{d}t+X_t^2\mathrm{d}B_t$.

12. 求 $\mathrm{d}X_t$，其中 $X_t=B_t^n$，$X_t=2+t+\mathrm{e}^{B_t}$，$X_t=[B_t^{(1)}]^2+[B_t^{(2)}]^2$，$(B_t^{(1)}, B_t^{(2)})$ 是二维 Brown 运动.

13. 记 $g_k(t)=E|B_t^k|$，证明 $g_k(t)=\frac{1}{2}k(k-1)\int_0^t g_{k-2}(s)\mathrm{d}s \quad (k\geqslant 2)$.

14. 设 $X_t=\mathrm{e}^{ct+\sum_{k=1}^{3}b_kB_t^{(k)}}$，其中 $(B_t^{(1)}, B_t^{(2)}, B_t^{(3)})$ 是三维 Brown 运动，求 $\mathrm{d}X_t$.

15. 计算 $\eta_t=\int_0^t B_s^3\mathrm{d}B_s$.

第 6 章　随机微分方程

6.1　随机微分方程

随机微分方程是描述相当广的一类物理现象、经济现象、金融现象的重要工具. 这些现象的共同点，是将输入归结为一个连续而变化极快的随机函数，也就是用一个 Brown 运动作为模型中的驱动力.

随机微分方程的一般形式为

$$\mathrm{d}\xi_t = b(t,\xi_t)\mathrm{d}t + \sigma(t,\xi_t)\mathrm{d}B_t, \tag{6.1}$$

其中 $b(t,x)$，$\sigma(t,x)$ 是给定的具有较为规则性质的函数. 这个随机微分方程应该理解成其积分形式，即

$$\xi_t = \xi_0 + \int_0^t b(s,\xi_s(\omega))\mathrm{d}s + \int_0^t \sigma(s,\xi_s)\mathrm{d}B_s, \tag{6.1$'$}$$

所以也称为**随机积分方程**，其中第一个积分理解为固定任意的 ω 后，对时间参数作普通函数的积分，而第二个积分即为 Ito 积分，而 ξ_0 是一个与 Brown 运动 $\{B_t: t\geqslant 0\}$ 相独立的随机初值.

6.1.1　非随机系数的随机微分方程

在一些工程问题中，会出现如下类型的随机微分方程

$$\xi_t^{(n)} + a_1\xi_t^{(n-1)} + \cdots + a_{n-1}\xi'_t = \sigma(t)\frac{\mathrm{d}B_t}{\mathrm{d}t},$$

其中 a_i 们都是常数，而 $\frac{\mathrm{d}B_t}{\mathrm{d}t}$ 是指 Brown 运动的形式导数（在第 5 章中，我们已经知道 Brown 运动的轨道函数没有导数，所以 $\frac{\mathrm{d}B_t}{\mathrm{d}t}$ 只是一个形式记号，表示它的积分是 Brown 运动，后者却是有确切含义的）.

在 $n=1$ 时，即

$$\mathrm{d}\xi_t = \sigma(t)\mathrm{d}B_t.$$

故有

$$\xi_t = \xi_0 + \int_0^t \sigma(s)\mathrm{d}B_s.$$

在 $n=2$ 时,即 $\mathrm{d}(\xi'_t+a\xi_t)=\sigma(t)\mathrm{d}B_t$,$\xi'_t+a\xi_t=\eta_0+\int_0^t\sigma(s)\mathrm{d}B_s$. 于是

$$\xi_t=\mathrm{e}^{-at}\left[\xi_0+\int_0^t\mathrm{e}^{au}\left(\eta_0+\int_0^u\sigma(s)\mathrm{d}B_s\mathrm{d}u\right)\right],\quad \eta_0=\xi'_0+a\xi_0.$$

这样可以逐步得到 n 阶情形时的表达式.

例 6.1 设 σ,b 都是正常数,η_0 是与 Brown 运动$\{B_t:t\geqslant 0\}$相互独立的正态随机变量. 显见

$$\eta_t\overset{\text{def}}{=\!=}\mathrm{e}^{-bt}\left(\eta_0+\sigma\int_0^t\mathrm{e}^{bs}\mathrm{d}B_s\right)$$

是 Gauss 过程. 它满足如下的随机微分方程

$$\mathrm{d}\eta_t=-b\eta_t\mathrm{d}t+\sigma\mathrm{d}B_t. \tag{6.2}$$

证明 令

$$\xi_t=\eta_0+\sigma\int_0^t\mathrm{e}^{bs}\mathrm{d}B_s,$$

则 $\eta_t=\mathrm{e}^{-bt}\xi_t$. 对它用 Ito 公式得到

$$\begin{aligned}\mathrm{d}\eta_t&=\mathrm{e}^{-bt}\mathrm{d}\xi_t+\xi_t\mathrm{d}\mathrm{e}^{-bt}+\mathrm{d}\mathrm{e}^{-bt}\mathrm{d}\xi_t\\&=\mathrm{e}^{-bt}\mathrm{d}\xi_t-b\mathrm{e}^{-bt}\xi_t\mathrm{d}t=\sigma\mathrm{d}B_t-b\eta_t\mathrm{d}t.\end{aligned}$$

这是一个随机微分方程,它最早出现于物理研究中,称为 **Langevin 方程**. 上面的推导正说明了 Langevin 方程(6.2)具有显式解

$$\eta_t=\mathrm{e}^{-bt}\left(\eta_0+\sigma\int_0^t\mathrm{e}^{bs}\mathrm{d}B_s\right).$$

因此,其平均值是

$$E\eta_t=E(\eta_0)\mathrm{e}^{-bt},$$

而其方差是

$$\begin{aligned}\mathrm{var}(\eta_t)&=\mathrm{var}(\eta_0)\mathrm{e}^{-2bt}+\sigma^2E\left[\int_0^t\mathrm{e}^{-b(t-s)}\mathrm{d}B_s\right]^2\\&=\mathrm{var}(\eta_0)\mathrm{e}^{-2bt}+\sigma^2\int_0^t\mathrm{e}^{-2b(t-s)}\mathrm{d}s\\&=\mathrm{var}(\eta_0)\mathrm{e}^{-2bt}+\frac{\sigma^2}{2b}(1-\mathrm{e}^{-2bt}).\end{aligned}$$

我们看到,对于任意方差有限的初值 η_0,当 $t\to+\infty$时恒有

$$E\eta_t\to 0,\quad \mathrm{var}(\eta_t)\to\frac{\sigma^2}{2b}.$$

这个显式解的积分部分 $\sigma\int_0^t\mathrm{e}^{-b(t-s)}\mathrm{d}B_s$ 是 Gauss 过程,而当 $t\to+\infty$ 时,初值 η_0 的影响越来越小,可见 η_t 具有正态的极限分布 $N\left(0,\frac{\sigma^2}{2b}\right)$. 它表示 Langevin 方程(其典型轨道的示意图可参见图 6.1) 的解在长时间运行后稳定在与初值无关的按正态分布的位置,其中方

差$\frac{\sigma^2}{2b}$表示随机分布的力 $\sigma \mathrm{d}B_t$ 与摩擦阻尼力 $-b\eta_t$ 之间的平衡的结果.

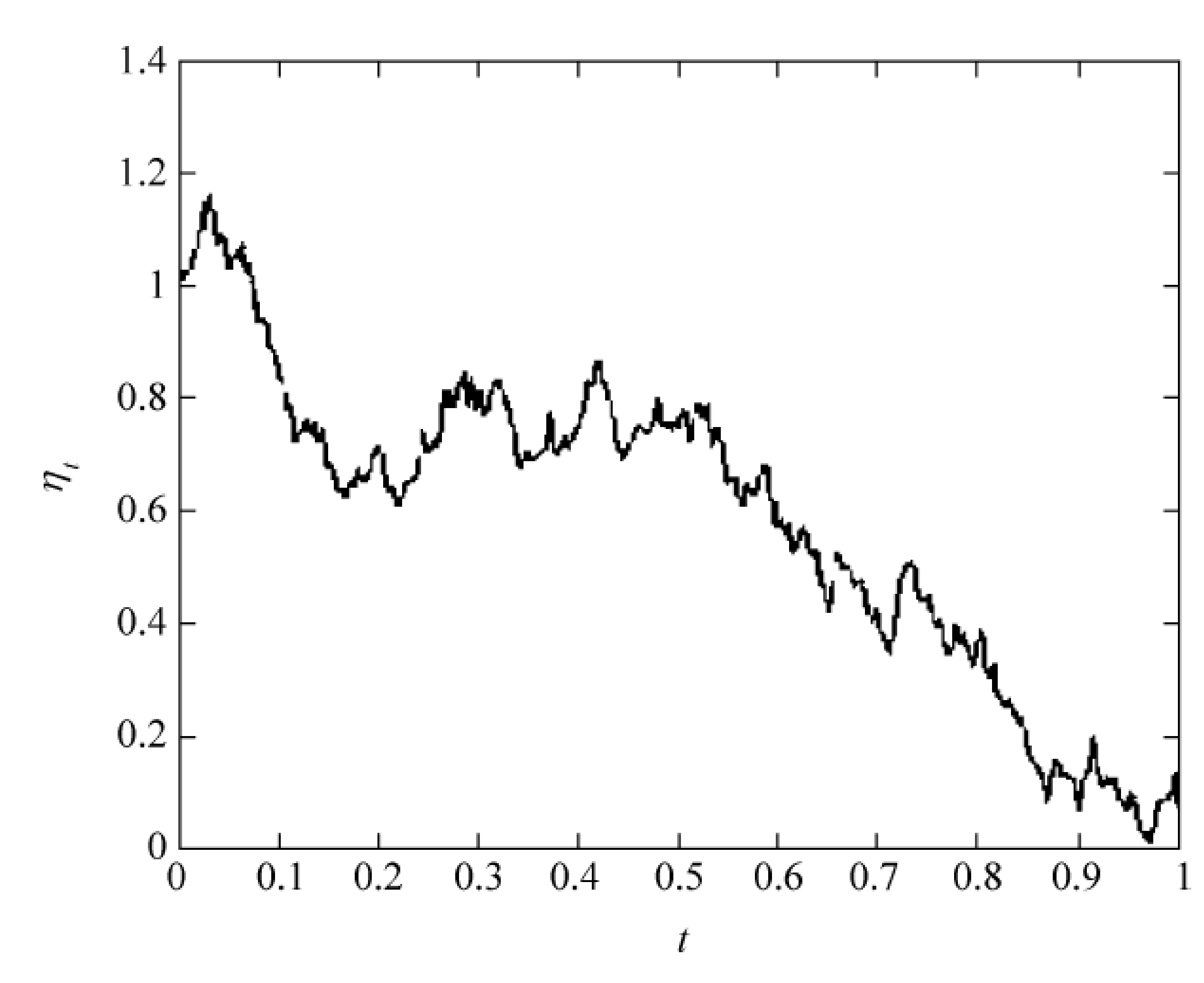


图 6.1　Langevin 方程轨道的示意图

我们还可以求得协方差函数 $\mathrm{cov}(\eta_t,\eta_s)$，并证明 $\lim\limits_{t-s\to 0}\mathrm{cov}(\eta_t,\eta_s)=0$.

Brown 运动的导数$\dot{B}_t=\frac{\mathrm{d}B_t}{\mathrm{d}t}$虽然处处不存在，但是我们仍然可以形式地将 $\int_a^b \Phi_t \dot{B}_t \mathrm{d}t$ 理解为$\int_a^b \Phi_t \mathrm{d}B_t$. 从而我们可以接受一个纯粹形式的记号$\dot{B}_t$，称为**白噪声**，它没有单独的意义，只是在出现于被积函数中时才起作用. 这种写法的优点是处理一些高阶的随机微分方程时较为直观，可以直接利用常微分方程的思维.

直接用 Brown 运动的数学模型描述生物学家 Brown 在液体中所观察看到的微小质点的随机运动并不十分理想. 一个更好的数学模型就是用 Langevin 方程描述这种运动的随机速度 v_t，即用

$$\mathrm{d}v_t=-bv_t\mathrm{d}t+\sigma \mathrm{d}B_t.$$

于是微小质点的随机运动是 $\xi_t=\xi_0+\int_0^t v_s \mathrm{d}s$，它形式地满足如下白噪声形式的 Ornstein-Uhlenbeck 方程

$$\begin{cases}\dfrac{\mathrm{d}^2\xi_t}{\mathrm{d}t^2}=-b\dfrac{\mathrm{d}\xi_t}{\mathrm{d}t}+\sigma\dot{B}_t,\\ \xi_0,\dfrac{\mathrm{d}\xi_t}{\mathrm{d}t}\Big|_{t=0}=v_0,\end{cases}$$

其中 ξ_0, v_0 都是独立的 Gauss 随机变量，并且与 Brown 运动$(B_t, t\geqslant 0)$独立，而 b 是摩擦系数，σ 是扩散系数.

考虑其对应的常微分方程

$$y''+by'=g.$$

它的中间积分是

$$y'+by=\int_0^t g(s)\mathrm{d}s+y'(0)+by(0).$$

由常数变异法得到解的公式为

$$y(t)=y(0)\mathrm{e}^{-bt}+\int_0^t \mathrm{e}^{-b(t-u)}\left[\int_0^u g(s)\mathrm{d}s+y'(0)+by(0)\right]\mathrm{d}u.$$

注意在随机的情形对应于 $\int_0^u g(s)\mathrm{d}s$ 的是 $\int_0^u \sigma\dot{B}_s\mathrm{d}s=\sigma B_u$，于是在随机情形的显式解是

$$\begin{aligned}\xi_t&=\xi_0\mathrm{e}^{-bt}+\int_0^t \mathrm{e}^{-b(t-u)}[\sigma B_u+v_0+b\xi_0]\mathrm{d}u\\&=\xi_0+\frac{1-\mathrm{e}^{-bt}}{b}v_0+\int_0^t\sigma\mathrm{e}^{-b(t-u)}B_u\mathrm{d}u.\end{aligned}$$

由此得到平均位移

$$E\xi_t=E\xi_0+\left(\frac{1-\mathrm{e}^{-bt}}{b}\right)Ev_0$$

和方差

$$\mathrm{var}(\xi_t)=\mathrm{var}(\xi_0)+\frac{(1-\mathrm{e}^{-bt})^2}{b^2}\mathrm{var}(v_0)+\sigma^2\int_0^t\int_0^t \mathrm{e}^{-b(t-u)}\mathrm{e}^{-b(t-s)}(u\wedge s)\mathrm{d}u\mathrm{d}s.$$

注意由 $\int_0^s \mathrm{e}^{bu}u\,\mathrm{d}u=s\frac{\mathrm{e}^{bs}}{b}-\frac{\mathrm{e}^{bs}-1}{b^2}$，得到

$$\begin{aligned}\int_0^t\int_0^t \mathrm{e}^{-b(t-u)}\mathrm{e}^{-b(t-s)}(u\wedge s)\mathrm{d}u\mathrm{d}s&=\int_0^t \mathrm{e}^{-2bt}\mathrm{e}^{bs}\left(\int_0^s \mathrm{e}^{bu}u\,\mathrm{d}u+s\int_s^t \mathrm{e}^{bu}\mathrm{d}u\right)\mathrm{d}s\\&=\mathrm{e}^{-2bt}\int_0^t \mathrm{e}^{bs}\left(\int_0^s \mathrm{e}^{bu}u\,\mathrm{d}u+s\frac{\mathrm{e}^{bt}-\mathrm{e}^{bs}}{b}\right)\mathrm{d}s\\&=\mathrm{e}^{-2bt}\int_0^t \mathrm{e}^{bs}\left(\frac{1-\mathrm{e}^{bs}}{b^2}+s\frac{\mathrm{e}^{bt}}{b}\right)\mathrm{d}s\\&=\mathrm{e}^{-2bt}\int_0^t\left(\frac{\mathrm{e}^{bs}-\mathrm{e}^{2bs}}{b^2}+s\mathrm{e}^{bs}\frac{\mathrm{e}^{bt}}{b}\right)\mathrm{d}s\\&=\mathrm{e}^{-2bt}\left[\frac{\mathrm{e}^{bt}-1}{b^3}-\frac{\mathrm{e}^{2bt}-1}{2b^3}+\frac{\mathrm{e}^{bt}}{b}\left(t\frac{\mathrm{e}^{bt}}{b}-\frac{\mathrm{e}^{bt}-1}{b^2}\right)\right]\\&=\frac{1}{2b^3}(4\mathrm{e}^{-bt}-\mathrm{e}^{-2bt}-3)+\frac{t}{b^2}.\end{aligned}$$

因此

$$\mathrm{var}(\xi_t)=\mathrm{var}(\xi_0)+\frac{(1-\mathrm{e}^{-bt})^2}{b^2}\mathrm{var}(v_0)+\frac{\sigma^2}{2b^3}(4\mathrm{e}^{-bt}-\mathrm{e}^{-2bt}-3)+\frac{\sigma^2 t}{b^2}.$$

用类似的推导，我们还可以得到协方差函数 $\mathrm{cov}(\xi_t,\xi_s)$ 的表达式，但是较为繁琐.

然而一般地,解的显式表示很难得到.所以得到平均位移的更可取的方法是直接利用随机微分方程来估计.在本例中我们可以通过积分关系与 Langevin 方程求得

$$E\xi_t = E\xi_0 + \int_0^t Ev_s \mathrm{d}s = E\xi_0 + \int_0^t \mathrm{e}^{-bs} Ev_0 \mathrm{d}s$$
$$= E\xi_0 + \left(\frac{1-\mathrm{e}^{-bt}}{b}\right)Ev_0,$$

其中用了取期望运算与求积分运算的交换性,通过一些数学的工具可以证明这样做是合理的.

例 6.2(随机调和振子)　形式随机微分方程

$$\begin{cases} \dfrac{\mathrm{d}^2\xi_t}{\mathrm{d}t^2} = -\lambda^2\xi_t + \dot{B}_t, \\ \xi_0, \dfrac{\mathrm{d}\xi_t}{\mathrm{d}t}\Big|_{t=0} = v_0. \end{cases}$$

对应的非齐次方程

$$y'' + \lambda^2 y = f$$

的通解为

$$y = \xi_0 \cos\lambda t + \frac{v_0}{\lambda}\sin\lambda t + \frac{1}{\lambda}\int_0^t \sin\lambda(t-s) f(s)\mathrm{d}s.$$

于是这个形式方程的解是

$$\xi_t = \xi_0 \cos\lambda t + \frac{v_0}{\lambda}\sin\lambda t + \frac{1}{\lambda}\int_0^t \sin\lambda(t-s)\mathrm{d}B_s,$$

其平均值出现为振动

$$E\xi_t = (E\xi_0)\cos\lambda t + \frac{Ev_0}{\lambda}\sin\lambda t,$$

而其方差为

$$\mathrm{var}(\xi_t) = \mathrm{var}(\xi_0)\cos^2\lambda t + \mathrm{var}\left(\frac{v_0}{\lambda}\right)\sin^2\lambda t + \frac{1}{\lambda^2}E\left[\int_0^t \sin\lambda(t-s)\mathrm{d}B_s\right]^2$$
$$= \mathrm{var}(\xi_0)\cos^2\lambda t + \mathrm{var}\left(\frac{v_0}{\lambda}\right)\sin^2\lambda t + \frac{1}{\lambda^2}\int_0^t \sin^2\lambda(t-s)\mathrm{d}s$$
$$= \mathrm{var}(\xi_0)\cos^2\lambda t + \mathrm{var}\left(\frac{v_0}{\lambda}\right)\sin^2\lambda t + \frac{1}{2\lambda^2}\left(t - \frac{\sin 2\lambda t}{2\lambda}\right).$$

用类似的推导,我们也可以得到协方差函数 $\mathrm{cov}(\xi_t, \xi_s)$ 的表达式.

6.1.2　随机系数的随机微分方程的例子

例 6.3　设 Φ_t 是有界的(即存在 $M>0$,使得对于任意 (t,ω) 有 $|\Phi_t(\omega)| \leqslant M$),且是 (B_t) 可知的随机过程. ς_t 是例 5.6 中的 Ito 积分的指数鞅 ς_t,即

$$\varsigma_t = \mathrm{e}^{\xi_t}, \quad \xi_t = \int_0^t \Phi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t \Phi_s^2 \mathrm{d}s.$$

那么,由 Ito 公式推出

$$d\varsigma_t = de^{\xi_t} = e^{\xi_t}d\xi_t + \frac{1}{2}e^{\xi_t}(d\xi_t)^2$$

$$= e^{\xi_t}\left(\Phi_t dB_t - \frac{1}{2}\Phi_t^2 dt\right) + \frac{1}{2}e^{\xi_t}\Phi_t^2 dt = e^{\xi_t}\Phi_t dB_t.$$

于是ς_t是随机微分方程

$$d\varsigma_t = \varsigma_t\Phi_t dB_t \tag{6.3}$$

满足初始条件

$$\varsigma_0 = 1$$

的解.事实上,由下面的定理 6.1 可以知道,它是这个方程的唯一解.

特别地,如果 $\Phi_t \equiv \lambda$,那么 $\eta_t = e^{\lambda B_t - \frac{1}{2}\lambda^2 t}$满足

$$d\eta_t = \lambda\eta_t dB_t. \qquad (6.3)'$$

例 6.4(常数 1 的多重不定限累次 Ito 积分) 将幂级数展开记为

$$e^{\lambda x - \frac{\lambda^2 t}{2}} = \sum_{n=0}^{+\infty}\lambda^n H_n(x,t),$$

其中 $H_n(x,t)(n \geqslant 0)$是 x 的(以 t 为参数)n 阶多项式,称为 ***n* 阶 Hermite 多项式**.

注意 $H_n(0,0)=0(n \geqslant 1)$.一般地,从展开式直接得到

$$H_0(x,t) = 1, \quad H_1(x,t) = x,$$

$$H_2(x,t) = \frac{x^2}{2} - \frac{t}{2}, \quad H_3(x,t) = \frac{x^3}{6} - \frac{tx}{2},$$

$$H_4(x,t) = \frac{x^4}{24} - \frac{tx^2}{4} + \frac{t^2}{8},$$

等等.于是对于例 6.3 中的 η_t 有

$$\eta_t = e^{\lambda B_t - \frac{1}{2}\lambda^2 t} = \sum_{n=0}^{+\infty}\lambda^n H_n(B_t,t).$$

将它代入方程(6.3)′,并比较 λ^n 的系数,得到

$$dH_{n+1}(B_t,t) = H_n(B_t,t)dB_t,$$

即

$$H_{n+1}(B_t,t) = H_{n+1}(B_0,0) + \int_0^t H_n(B_{t_n},t_n)dB_{t_n} = \int_0^t H_n(B_{t_n},t_n)dB_{t_n}.$$

因为 $H_1(x,t) = x$,所以 $H_1(B_t,t) = B_t = \int_0^t dB_{t_1}$.故而

$$H_2(B_t,t) = \int_0^t H_1(B_{t_1},t_1)dB_{t_1} = \int_0^t\int_0^{t_1}dB_{t_2}dB_{t_1}.$$

因此,常数 1 的二重不定限累次 Ito 积分是

$$\int_0^t\int_0^{t_1} dB_{t_2} dB_{t_1} = H_2(B_t,t) = \frac{1}{2}B_t^2 - \frac{t}{2}.$$

常数1的三重不定限累次 Ito 积分是

$$\int_0^t\int_0^{t_1}\int_0^{t_2} dB_{t_3} dB_{t_2} dB_{t_1} = H_3(B_t,t) = \frac{1}{6}B_1^3 - \frac{1}{2}tB_t.$$

由数学归纳法容易证明常数1的 n 重不定限累次 Ito 积分为 $H_n(B_t,t)$，即

$$\int_0^t\int_0^{t_1}\cdots\int_0^{t_{n-1}} dB_{t_n} dB_{t_{n-1}}\cdots dB_{t_1} = \int_0^t H_{n-1}(B_{t_{n-1}},t_{n-1})dB_{t_{n-1}} = H_n(B_t,t).$$

注意，如果将 Brown 运动 B_t 换为普通的自变量 t，那么得到的是常数1的 n 重不定限累次积分，其值为 $\frac{1}{n!}t^n$. 两者比较，在多重不定限累次 Ito 积分的表达式中只是以 $H_n(B_t,t)$ 代替了 $\frac{1}{n!}t^n$.

下面我们推导 Hermite 多项式 $H_n(x,t)(n\geqslant 0)$ 的显式表示. 由于

$$\frac{d^n}{d\lambda^n}(e^{-\frac{(x-\lambda t)^2}{2t}}) = (-t)^n\left[\frac{d^n}{du^n}e^{-\frac{u^2}{2t}}\right]_{u=x-\lambda t},$$

故

$$\left[\frac{d^n}{d\lambda^n}(e^{-\frac{(x-\lambda t)^2}{2t}})\right]_{\lambda=0} = (-t)^n\left[\frac{d^n}{dx^n}e^{-\frac{x^2}{2t}}\right].$$

于是

$$\left[\frac{d^n}{d\lambda^n}(e^{\lambda x-\frac{\lambda^2 t}{2}})\right]_{\lambda=0} = (-t)^n e^{x^2/2t}\frac{d^n}{dx^n}e^{-\frac{x^2}{2t}}.$$

所以由 Taylor 展开式，我们得到 Hermite 多项式 $H_n(x,t)(n\geqslant 0)$ 的显式表示为

$$H_n(x,t) = \frac{1}{n!}\left[\frac{d^n}{d\lambda^n}(e^{\lambda x-\frac{1}{2}\lambda^2 t})\right]_{\lambda=0} = \frac{(-t)^n}{n!}e^{\frac{x^2}{2t}}\frac{d^n}{dx^n}e^{-\frac{x^2}{2t}}.$$

例 6.5（描述证券价格的 Black-Scholes 模型）　Black-Scholes 用如下的线性随机微分方程

$$d\xi_t = \xi_t(\mu dt + \sigma dB_t) \tag{6.4}$$

的解 ξ_t 来描述证券价格随时间随机变化的数学模型. 这个方程称为 **Black-Scholes 随机微分方程**，或 **Black-Scholes 模型**，其中常数 μ 称为(平均)**收益率**(yield)，常数 σ 称为**波动率**(volatility).

我们可以用 Ito 公式直接验证

$$\xi_t \stackrel{\text{def}}{=} \xi_0 e^{\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma B_t}$$

是 Black-Scholes 随机微分方程的解. 当 $\xi_0=1$ 时，它就是几何 Brown 运动，其轨道的示意图可参见图 6.2.

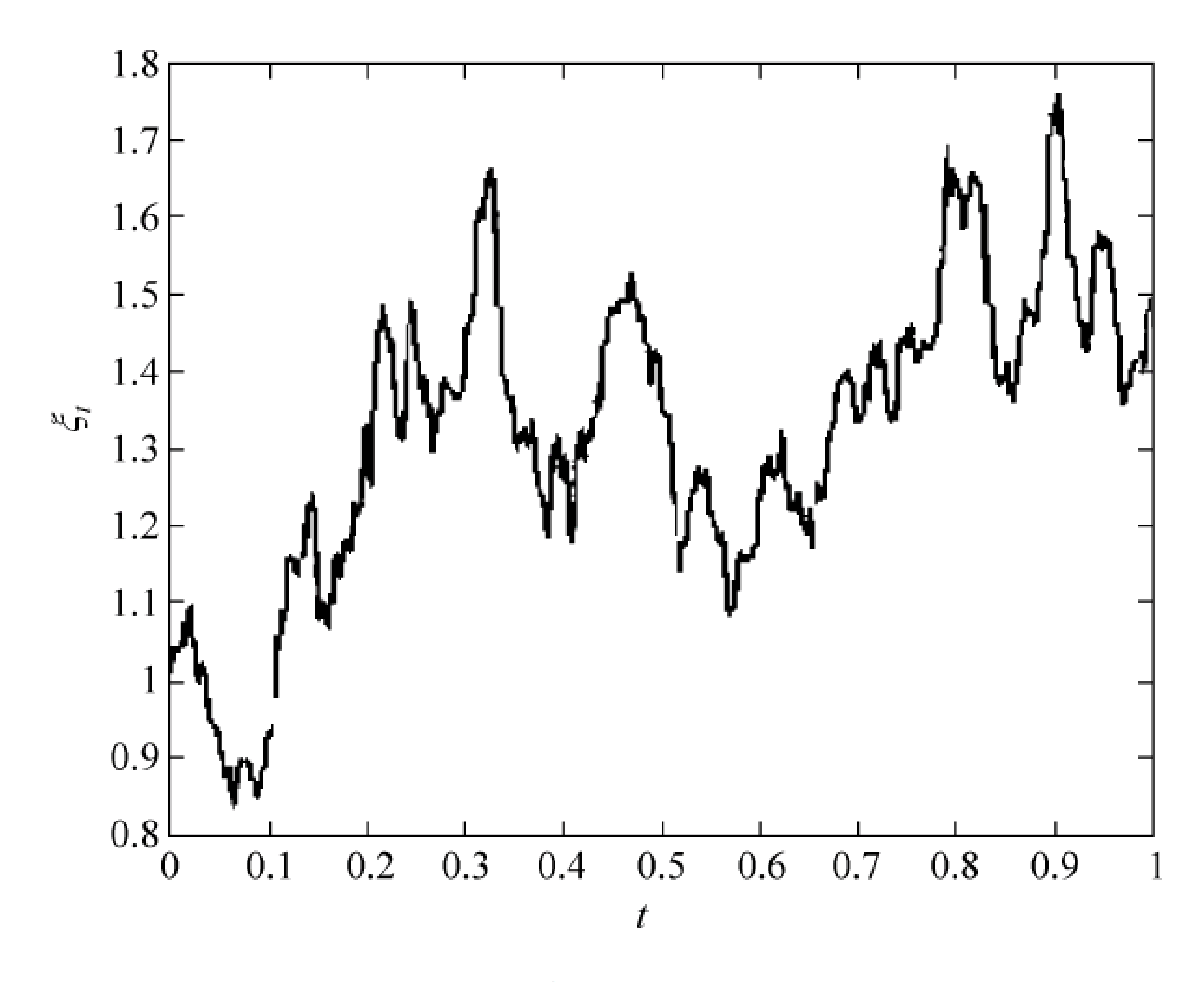


图 6.2　证券价格变化的图示

6.1.3　随机微分方程的解的存在唯一性

随机微分方程解的存在性和唯一性的研究，是随机微分方程的理论研究的一个基本问题. 它是随机分析的理论基础，对应用问题也有重要意义. 由正确的物理规律得到的随机微分方程，应该存在唯一的解. 反之，如果在一些领域中建立的随机微分方程的解并不存在，那么，很可能在建模时由于忽略了一些不能忽视的因素而导致了错误的模型. 又如果从不同来源信息得到的是同一个随机微分方程，而且此方程有唯一的解，这说明，从这些不同来源得到的信息有相同的数学规律.

我们给出随机微分方程的一个最简单的存在唯一性定理.

定理 6.1　如果 $\sigma(t,x)$, $b(t,x)$ 满足：对于任意 $T>0$，存在 $C_T>0$，使得对于任意 $t\leqslant T$，一致地满足如下的 **Lipschitz 条件**

$$|\sigma(t,x)-\sigma(t,y)|+|b(t,x)-b(t,y)|\leqslant C_T|x-y|.$$

而具有有限方差的初始随机变量 ξ_0 与 Brown 运动 $\{B_t, t\geqslant 0\}$ 独立. 那么，在 ξ_0 给定的条件下，随机微分方程(6.1)存在唯一的解 ξ_t，其含义为：如果有另一个解 ξ_t'，那么只要满足 $P(\xi_0'=\xi_0)=1$，就一定有 $P(\xi_t'=\xi_t, \forall t>0)=1$. 它可以由下面的迭代程序的极限得到

$$\xi_t^{(0)} = \xi_0,$$

$$\xi_t^{(n+1)} = \xi_0 + \int_0^t b(s,\xi_s^{(n)})\mathrm{d}s + \int_0^t \sigma(s,\xi_s^{(n)})\mathrm{d}B_s.$$

即对任意 $T>0$ 有

$$E(\sup_{t\leqslant T}|\xi_t^{(n)}-\xi_t|^2)\xrightarrow{n\to+\infty}0.$$

它也蕴含了：对任意 $\varepsilon>0,T>0$ 有

$$P(\sup_{t\leqslant T}|\xi_t^{(n)}-\xi_t|>\varepsilon)\xrightarrow{n\to+\infty}0.$$

此定理中的迭代近似随机过程序列，正好是随机微分方程的解的一个近似算法.

推论 6.2　随机微分方程的解作为随机过程 $\{\xi_t:t\geqslant0\}$ 是 (B_t) 可知的，且是初始随机变量 ξ_0 与 Brown 运动在 t 前(随机)历史

$$(B_s)_{0\leqslant s\leqslant t}=\{B_u:u\leqslant t\}$$

的某个"泛函"关系(粗略地说，依赖于一段连续历史的一个对应关系，称为这段连续历史的一个**泛函**(functional)关系)：

$$\xi_t=G(t,\xi_0,(B_s)_{0\leqslant s\leqslant t}).$$

直观证明　从极限列的构造利用归纳法易见，随机微分方程的解的迭代过程的每一步，即近似随机过程 $\{\xi_t^{(n)}:t\geqslant0\}$ 都是 (B_t) 可知的，也就是在时刻 t 前都可以写成初始随机变量 ξ_0 与 Brown 运动在 t 前历史的泛函关系，即

$$\xi_t^{(n)}=G_n(t,\xi_0,(B_s)_{0\leqslant s\leqslant t}).$$

于是随机微分方程的解 $\{\xi_t:t\geqslant0\}$ 作为其极限，就可以写成

$$\xi_t=G(t,\xi_0,(B_s)_{0\leqslant s\leqslant t}),$$

所以作为随机过程是 (B_t) 可知的.

注 1　保证随机微分方程的解的存在唯一性的系数条件还可以减弱. 例如，可以取消 ξ_0 有有限方差的限制，并可以将系数的一致 Lipschitz 条件假定减弱为

(1) **局部一致 Lipschitz 条件**，即对于任意 N，存在常数 C_N，使得对于任意 $|x|,|y|\leqslant N$，以及任意 T，对于 $t\leqslant T$，都有

$$|\sigma(t,x)-\sigma(t,y)|+|b(t,x)-b(t,y)|\leqslant C_N|x-y|.$$

(2) **线性增长条件**，即任意 T，存在常数 C_T，对于 $t\leqslant T$ 满足

$$|\sigma(t,x)|+|b(t,x)|\leqslant C_T(1+|x|).$$

再假定初始随机变量 ξ_0 与 Brown 运动 $\{B_t,t\geqslant0\}$ 独立，那么，在 ξ_0 给定的条件下，随机微分方程(6.1)在 $0\leqslant t<+\infty$ 上存在唯一的解 ξ_t(这时它的方差可能为 $+\infty$). 并且这个解可以由下面方式近似地得到：令

$$\sigma_n(x)\stackrel{\text{def}}{=\!=}\begin{cases}\sigma(t,x), & |x|\leqslant n,\\ \sigma(t,n), & x>n,\\ \sigma(t,-n), & x<-n,\end{cases}\qquad b_n(x)\stackrel{\text{def}}{=\!=}\begin{cases}b(t,x) & |x|\leqslant n,\\ b(t,n) & x>n,\\ b(t,-n) & x<-n,\end{cases}$$

$$\xi_t^{(n+1)}=\xi_0+\int_0^t b_n(s,\xi_s^{(n)})\mathrm{d}s+\int_0^t\sigma_n(s,\xi_s^{(n)})\mathrm{d}B_s.$$

这里"近似"的含义是：对任意 $\varepsilon>0,T>0$ 恒有

$$P(\sup_{t\leqslant T}|\xi_t^{(n)}-\xi_t|>\varepsilon)\xrightarrow{n\to+\infty}0.$$

这种收敛在概率论中称为**局部一致地按概率收敛**.

如果线性增长条件不满足,那么,经过一些处理仍然可以得到随机微分方程存在唯一的一个在扩展意义下的解,它可能在某个有限时刻(当然此时刻依赖与轨道,因此是随机时刻)ζ 及以后恒取值 $-\infty$ 或 $+\infty$. 这个随机时刻称为**爆炸时刻**,这种解称为**有爆炸的解**. 这时,在任意给定的时刻 t 处,随机变量 ξ_t 就有可能以一定的概率取值 $-\infty$ 或 $+\infty$. 这是与普通的随机变量的不同之处. 如果将 $-\infty$ 和 $+\infty$ 作为"附加值"加进实数集 $\mathbb{R}$(称为两点"紧化")记为 $\overline{\mathbb{R}}$,则这个解可以视为取值于 $\overline{\mathbb{R}}$ 的随机过程. 但是这个过程的轨道一旦到达 $-\infty$ 或 $+\infty$ 后就永远停留在那里,即 $-\infty$ 和 $+\infty$ 都是过程的吸收状态,或者称之为"坟墓".

注 2 还有一种随机微分方程,其系数不仅依赖于 t 和 ξ_t,还依赖于解的过去的整个历史 $\{\xi_s, s\leqslant t\}$,这种随机微分方程称为**泛函随机微分方程**. 这时的情况更为复杂.

例 6.1、例 6.2 与例 6.5 中的随机微分方程,都满足定理 6.1 的条件. 所以在初始值给定的条件下,方程的解是存在唯一的. 而实际上我们已经得到了其解的显式表示. 但是,在例 6.3 中的随机微分方程,却不满足定理 6.1 的条件. 事实上,它的解是唯一的,然而需要另辟途经证明,本书从略.

定义 6.1 如果随机微分方程的系数不显含 t,即

$$\sigma(t,x)=\sigma(x), \quad b(t,x)=b(x),$$

则称为**时齐的**随机微分方程.

在通常的随机建模中,更多地遇见的是时齐的随机微分方程.

在图 6.3 中的上图与下图分别是常微分方程

$$\begin{cases} \mathrm{d}x_t = b(x_t)\mathrm{d}t, \\ x_0 = x \end{cases}$$

的解与随机微分方程

$$\begin{cases} \mathrm{d}\xi_t = b(\xi_t)\mathrm{d}t + \sigma(\xi_t)\mathrm{d}B_t, \\ \xi_0 = x \end{cases}$$

的解的样本的示意比较图.

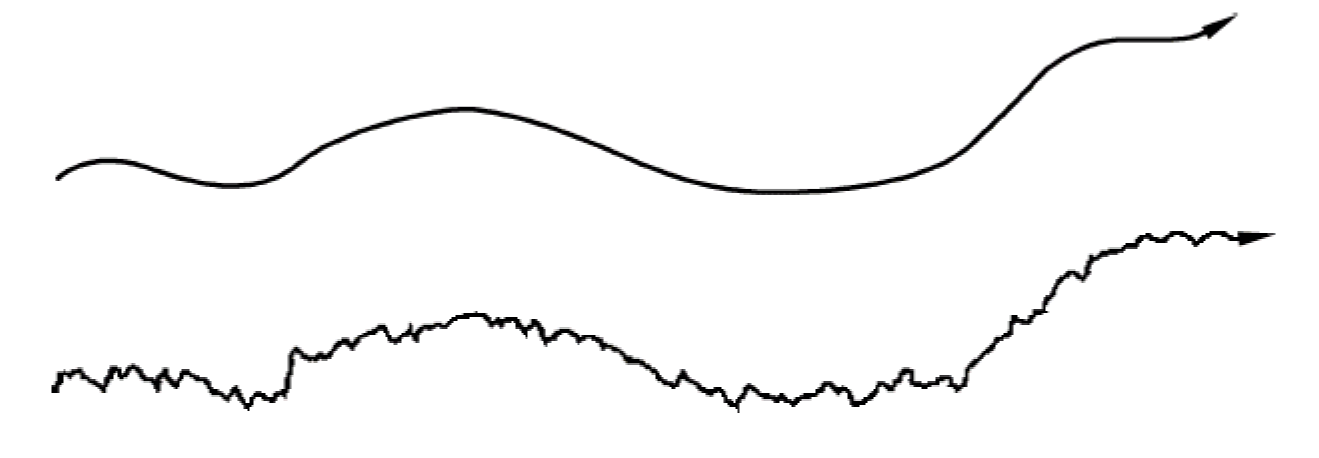


图 6.3 常微分方程与随机微分方程的解的比较示意图

例 6.6 随机微分方程

$$d\xi_t = F(t)\xi_t dt + C(t)dB_t,$$

其中 $F(t)$,$C(t)$是非随机的数值函数,假定它们在任意有限的时间区间上有界.我们可以仿照常微分方程的方法求解.首先将方程改写为

$$d\xi_t - F(t)\xi_t dt = C(t)dB_t. \tag{6.5}$$

它对应的齐次线性随机微分方程为

$$d\xi_t - F(t)\xi_t dt = 0.$$

此方程的解是

$$\xi_t = \xi_0 e^{\int_0^t F(s)ds}.$$

再仿照常微分方程中的恰当因子方法,将 $e^{-\int_0^t F(s)ds}$ 乘到(6.5)式的两边,得到

$$d(\xi_t e^{-\int_0^t F(s)ds}) = C(t)e^{-\int_0^t F(s)ds}dB_t.$$

于是有

$$\xi_t = \xi_0 e^{\int_0^t F(s)ds} + \int_0^t C(u)e^{\int_u^t F(s)ds}dB_u. \tag{6.6}$$

例 6.7(线性随机微分方程) 一般的线性随机微分方程为

$$d\xi_t = (b\xi_t + a)dt + (\sigma\xi_t + c)dB_t, \tag{6.7}$$

其中的 b,a,σ,c 都是常数.

我们可以仿照常微分方程求解的方法,得到它的解.首先,将方程改写为

$$d\xi_t - (b\xi_t dt + \sigma\xi_t dB_t) = a dt + c dB_t.$$

它所对应的齐次线性随机微分方程为

$$d\eta_t - \eta_t(b dt + \sigma dB_t) = 0,$$

它正是一个 Black-Scholes 随机微分方程.此方程的解是唯一地存在的,由例 6.3 知道其解是

$$\eta_t = \eta_0 e^{\left(b-\frac{\sigma^2}{2}\right)t+\sigma B_t}.$$

再仿照常微分方程中的恰当因子方法,将其倒数 η_t^{-1} 乘到非齐次方程上,以使用 Ito 公式求得 $d(\xi_t\eta_t^{-1})$.由

$$\eta_t^{-1} = \eta_0^{-1}e^{-\left(b-\frac{\sigma^2}{2}\right)t-\sigma B_t} = \eta_0^{-1}e^{\left[(-b+\sigma^2)-\frac{\sigma^2}{2}\right]t-\sigma B_t},$$

可见 η_t^{-1} 应该满足以下方程

$$d\eta_t^{-1} = \eta^{-1}[(-b+\sigma^2)dt - \sigma dB_t].$$

再用 Ito 公式得到

$$\begin{aligned}
d(\xi_t\eta_t^{-1}) &= \xi_t d\eta_t^{-1} + \eta_t^{-1}d\xi_t + d\xi_t d\eta_t^{-1} \\
&= \eta_t^{-1}\xi_t[(-b+\sigma^2)dt - \sigma dB_t] + \eta_t^{-1}[(b\xi_t + a)dt + (\sigma\xi_t + c)dB] - (\sigma\xi_t + c)\eta_t^{-1}\sigma dt \\
&= \eta_t^{-1}[(a - c\sigma)dt + c dB].
\end{aligned}$$

由于我们已经求得了 η_t 的表达式，故而我们就得到解 ξ_t 的显式表示

$$\begin{aligned}\xi_t &= \eta_t\left[\xi_0\eta_0^{-1} + (a - c\sigma)\int_0^t \eta_s^{-1}\mathrm{d}s + c\int_0^t \eta_s^{-1}\mathrm{d}B_s\right] \\ &= \xi_0 \mathrm{e}^{\left(b-\frac{\sigma^2}{2}\right)t+\sigma B_t} + (a - c\sigma)\int_0^t \mathrm{e}^{\left(b-\frac{\sigma^2}{2}\right)(t-s)+\sigma(B_t-B_s)}\mathrm{d}s + c\int_0^t \mathrm{e}^{\left(b-\frac{\sigma^2}{2}\right)(t-s)+\sigma(B_t-B_s)}\mathrm{d}B_s. \quad (6.8)\end{aligned}$$

我们还可以通过 Brown 运动的期望和方差求得解 ξ_t 的期望、方差和协方差.

例 6.8（时变系数的线性随机微分方程） 对于随机微分方程

$$\mathrm{d}\xi_t = (a(t) + b(t)\xi_t)\mathrm{d}t + (c(t) + \sigma(t)\xi_t)\mathrm{d}B_t,$$

记其相应的齐次方程

$$\mathrm{d}\xi_t^{(0)} = b(t)\xi_t^{(0)}\mathrm{d}t + \sigma(t)\xi_t^{(0)}\mathrm{d}B_t$$

的解为 $\xi_t^{(0)}$. 我们再次仿照常微分方程中的常数变异法待定 $C(t)$，确定非齐次方程的形如

$$\xi_t = C(t)\xi_t^{(0)}$$

的解. 对此利用 Ito 公式得到

$$\begin{aligned}\mathrm{d}\xi_t &= C(t)\mathrm{d}\xi_t^{(0)} + \xi_t^{(0)}C'(t)\mathrm{d}t \\ &= [b(t)\xi_t\mathrm{d}t + \sigma(t)\xi_t\mathrm{d}B_t] + \xi_t^{(0)}C'(t)\mathrm{d}t.\end{aligned}$$

因此必须选取 $C(t)$ 满足

$$\xi_t^{(0)}C'(t) = a(t)\mathrm{d}t + c(t)\mathrm{d}B_t$$

即取

$$C(t) = C + \int_0^t \xi_s^{(0)-1}(a(s)\mathrm{d}s + c(s)\mathrm{d}B_s)$$

即可.

另一方面，我们也可以将例 6.7 中得到的公式，直接推广到系数显含时间 t 的情形. 即

$$\mathrm{d}\xi_t = (b(t)\xi_t + a(t))\mathrm{d}t + (\sigma(t)\xi_t + c(t))\mathrm{d}B_t, \quad (6.7)'$$

用类似的考虑得到

$$\begin{aligned}\xi_t =& \xi_0 \mathrm{e}^{\int_0^t\left(b(s)-\frac{\sigma^2(s)}{2}\right)\mathrm{d}s+\int_0^t\sigma(s)\mathrm{d}B_s} + \int_0^t (a(s) - c(s)\sigma(s))\mathrm{e}^{\int_s^t\left(b(u)-\frac{\sigma^2(u)}{2}\right)\mathrm{d}u+\int_s^t\sigma(u)\mathrm{d}B_u}\mathrm{d}s \\ &+ \int_0^t c(s)\mathrm{e}^{\int_s^t\left(b(u)-\frac{\sigma^2(u)}{2}\right)\mathrm{d}u+\int_s^t\sigma(u)\mathrm{d}B_u}\mathrm{d}B_s. \quad (6.8)'\end{aligned}$$

对于 m 维 Brown 运动 $\boldsymbol{B}_t$，即 m 个相互独立的 Brown 运动排成的列向量. 对于 $d\times m$ 矩阵值函数 $\boldsymbol{\Sigma}(\boldsymbol{x})$ 及 d 维向量值函数 $\boldsymbol{b}(\boldsymbol{x})$，考虑 d 维随机微分方程

$$\boldsymbol{\xi}_t = \boldsymbol{x} + \int_0^t \boldsymbol{b}(\boldsymbol{\xi}_s)\mathrm{d}s + \int_0^t \boldsymbol{\Sigma}(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s.$$

它也有与定理 6.1 类似的存在唯一性条件. 也有与例 6.1～例 6.3 中相应的方程，以及其解的表达式.

例 6.9（Stratonovich 型随机微分方程） 随机微分方程

$$\mathrm{d}\xi_t = b(t)\xi_t\mathrm{d}t + \sigma(t)\xi_t \circ \mathrm{d}B_t$$

的解是

$$\xi_t = \xi_0 e^{\int_0^t b(s)\mathrm{d}s + \int_0^t \sigma(s)\circ \mathrm{d}B_s}.$$

这可以通过 Ito 公式直接验证.

对于一个 Ito 随机微分方程,如果将它形式地视为由白噪声驱动的微分方程,并且用光滑的随机过程列近似白噪声$\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$(例如随机差商的光滑近似),那么,在理论上可以证明,相应的带有随机参数的常微分方程的解并不依概率趋于这个 Ito 方程的解,而是依概率趋于将 Ito 积分$\int_0^t \sigma(\xi_t)\mathrm{d}B_t$用 Stratonovich 积分$\int_0^t \sigma(\xi_t)\circ \mathrm{d}B_t$代替后的 Stratonovich 随机微分方程的解. 这就说明了用 Ito 随机微分方程建模会失去光滑逼近时的稳定性. 而用 Stratonovich 型随机微分方程建模对白噪声的光滑的随机逼近,则其解有稳定性. 所以,有人认为从建模的角度看,用 Stratonovich 型随机微分方程建模更为合理.

例 6.10(Brown 桥)　对随机过程

$$X_t = (1-t)\int_0^t \frac{1}{1-s}\mathrm{d}B_s$$

用 Ito 公式,立刻得到它满足随机微分方程

$$\mathrm{d}X_t = -\frac{X_t}{1-t}\mathrm{d}t + \mathrm{d}B_t \quad (0 \leqslant t < 1),$$
$$X_0 = 0.$$

这个随机过程具有性质$\lim\limits_{t\to 1} X_t = 0$,称为在时间 0 与时间 1 间的 **Brown 桥**(见图 6.4)(注意在考察这个方程的解的存在唯一性时,这里的时间 1 要当作$+\infty$来处理).

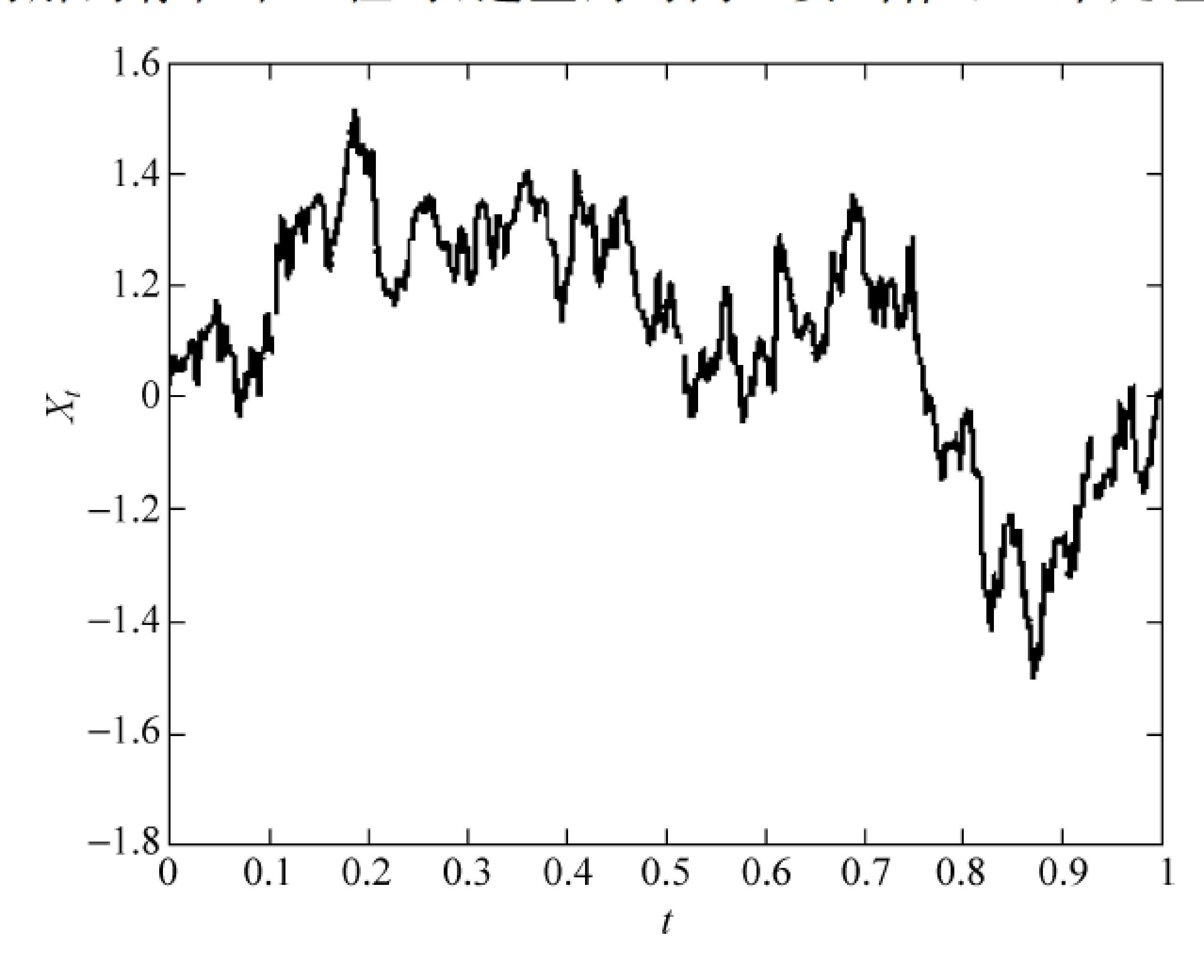


图 6.4　Brown 桥的轨道的示意图

例 6.11 考虑随机微分方程

$$d\xi_t = (a\xi_t + b\sqrt{\xi_t})dt + c\xi_t dB_t.$$

如果将 $at+cB_t$ 看成"自变量",那么这个随机微分方程很像常微分方程中的 Bernoulli 方程.它来自金融中的随机波动率模型.我们可以通过变换,并运用 Ito 公式,把此方程化为一个可以将随机性的样本点 ω 看成为参变量的普通的常微分方程,从而可以求得方程的解的表达式

$$\xi_t = e^{cB_t - \frac{c^2}{2}t + at}\left(\frac{b}{2}\int_0^t e^{-\frac{c}{2}B_s - \left(\frac{c^2}{4}+\frac{a}{2}\right)s} ds + \sqrt{\xi_0}\right)^2.$$

以下是此公式的推导过程.令

$$M_t = e^{-cB_t - \left(\frac{1}{2}c^2 + a\right)t},$$

由例 6.3,我们有

$$dM_t = -M_t(c dB_t + a dt).$$

再用 Ito 公式得到

$$\begin{aligned} d(M_t\xi_t) &= M_t d\xi_t + \xi_t dM_t + dM_t d\xi_t \\ &= M_t[(a\xi_t + b\sqrt{\xi_t})dt + c\xi_t dB_t] - \xi_t M_t(c dB_t + a dt) + (-cM_t)(c\xi_t)dt \\ &= M_t\xi_t(-c^2 dt) + b\sqrt{M_t}\sqrt{M_t\xi_t}\,dt. \end{aligned}$$

让我们置

$$\eta_t = M_t\xi_t \quad (\text{注意 } \eta_0 = \xi_0),$$

那么 η_t 满足如下的随机系数的一阶常微分方程(Bernoulli 方程)

$$\frac{d\eta_t}{dt} = -c^2\eta_t + b\sqrt{M_t}\sqrt{\eta_t},$$

由复合微分公式:

$$d\sqrt{\eta_t} = \frac{1}{2\sqrt{\eta_t}} d\eta_t.$$

于是 η_t 的方程就等价于

$$\frac{d\sqrt{\eta_t}}{dt} = -\frac{c^2}{2}\sqrt{\eta_t} + \frac{b}{2}\sqrt{M_t}.$$

作变换

$$\zeta_t = \sqrt{\eta_t} \quad (\zeta_0 = \sqrt{\xi_0}),$$

就得到 ζ_t 的具有随机的非齐次项的一阶常系数线性常微分方程

$$\frac{d\zeta_t}{dt} + \frac{c^2}{2}\zeta_t = \frac{b}{2}e^{-\frac{1}{2}cB_t - \left(\frac{1}{4}c^2 + \frac{a}{2}\right)t}.$$

即

$$(\zeta_t e^{\frac{c^2}{2}t})' = \frac{b}{2} e^{-\frac{1}{2}cB_t - (\frac{1}{4}c^2 + \frac{a}{2})t}.$$

则

$$\zeta_t = \left[\zeta_0 + \frac{b}{2}\int_0^t e^{-\frac{1}{2}cB_{ts} - (\frac{1}{4}c^2 + \frac{a}{2})s} ds\right] e^{-\frac{c^2}{2}t}.$$

代回去便得到显式解

$$\xi_t = \frac{\zeta_t^2}{M_t} = e^{cB_t + (a - \frac{1}{2}c^2)t} \left[\sqrt{\xi_0} + \frac{b}{2}\int_0^t e^{-\frac{1}{2}cB_{ts} - (\frac{1}{4}c^2 + \frac{a}{2})s} ds\right]^2.$$

类似的思维,可以用来求解 $d\xi_t = (a\xi_t + b\xi_t^\alpha)dt + c\xi_t dB_t$,其中 α 是一个实数.

像这样可以得到较为简单的显式解的随机微分方程并不多见,绝大多数的随机微分方程的解不可能有简单的表达式. 因此,解的性质的讨论,以及求近似解,就显得更为重要.

例 6.12(多维情形) 设 $\boldsymbol{D}(t)$是 d 阶决定性矩阵,$\boldsymbol{c}(t)$,$\boldsymbol{\sigma}(t)$是 d 维决定性向量. 形式随机微分方程

$$\frac{d\boldsymbol{\xi}_t}{dt} = \boldsymbol{c}(t) + \boldsymbol{D}(t)\boldsymbol{\xi}_t + \boldsymbol{\sigma}(t)\frac{d\boldsymbol{B}_t}{dt}$$

的解,可以由类似地采用常数变异法求得,具体为

$$\boldsymbol{\xi}_t = \boldsymbol{F}(t)\left[\boldsymbol{\xi}_0 + \int_0^t F(s)^{-1}(\boldsymbol{c}(s)ds + \boldsymbol{\sigma}(s)d\boldsymbol{B}_s)\right],$$

其中 $\boldsymbol{F}(t)$是非自治线性常微分方程

$$\frac{d\boldsymbol{x}_t}{dt} = \boldsymbol{D}(t)\boldsymbol{x}_t$$

的基本解,即它是满足

$$\begin{cases} \dfrac{d\boldsymbol{F}(t)}{dt} = \boldsymbol{D}(t)\boldsymbol{F}(t), \\ \boldsymbol{F}(0) = \boldsymbol{I} \end{cases}$$

的唯一解.

引理 6.3 Ito 方程与 Stratonovich 型随机微分方程间有如下的转换公式

$$d\boldsymbol{\xi}_t = \boldsymbol{b}(t,\boldsymbol{\xi}_t)dt + \boldsymbol{\Sigma}(t,\boldsymbol{\xi}_t)d\boldsymbol{B}_t$$

等价于

$$d\boldsymbol{\xi}_t = \left[\boldsymbol{b}(t,\boldsymbol{\xi}_t) - \frac{1}{2}\boldsymbol{c}(t,\boldsymbol{\xi}_t)\right]dt + \boldsymbol{\Sigma}(t,\boldsymbol{\xi}_t)\circ d\boldsymbol{B}_t,$$

其中

$$\boldsymbol{c}(t,\boldsymbol{x}) = \begin{pmatrix} c_1(t,\boldsymbol{x}) \\ c_2(t,\boldsymbol{x}) \\ \vdots \\ c_d(t,\boldsymbol{x}) \end{pmatrix}, \quad c_i(t,\boldsymbol{x}) = \sum_{j=1}^d \sum_{k=1}^m \sigma_{jk}(t,\boldsymbol{x})\frac{\partial \sigma_{ik}(t,\boldsymbol{x})}{\partial x_j}, \quad \boldsymbol{\Sigma}(t,\boldsymbol{x}) = (\sigma_{ik}(t,\boldsymbol{x})).$$

特别地

$$d\xi_t = b(t,\xi_t)dt + \sigma(t,\xi_t)dB_t$$

等价于

$$d\xi_t = \left[b(t,\xi_t) - \frac{1}{2}\sigma'(t,\xi_t)\sigma(t,\xi_t)\right]dt + \sigma(t,\xi_t)\circ dB_t.$$

一维随机微分方程是最简单的情形，结论也相对地多一些. 对于时齐的一维随机微分方程，只要假定系数有一定的可微性，就可以通过解两个常微分方程得到随机微分方程的解. 这将在 6.2 节中介绍.

*6.2 通过两个常微分方程的解给出光滑系数的一维随机微分方程的解

6.2.1 Ito 公式的简单推广

命题 6.4（推广的 Ito 公式） 设 B_t 为 Brown 运动，而 $\{Y_t, t\geqslant 0\}$ 是一个 (B_t) 可知的，而且其样本函数 $Y_t(\omega)$ 是对 t 连续可微的随机过程. 又实函数 $f(y,x)$ 对 x 二次连续可微，且对 y 连续可微. 那么，随机过程 $\eta_t = f(Y_t, B_t)$ 也是 Ito 过程，而且有

$$\begin{aligned} d\eta_t &= df(Y_t,B_t) + \frac{1}{2}d^2 f(Y_t,B_t) \\ &= f'_y(Y_t,B_t)\frac{dY_t}{dt}dt + f'_x(Y_t,B_t)dB_t + \frac{1}{2}f''_{xx}(Y_t,B_t)dt. \end{aligned}$$

下面我们将说明，只要系数足够光滑（即足够多次连续可微），随机微分方程的解就可以用两个常微分方程的解表达.

6.2.2 通过两个常微分方程的解求解 Stratonovich 随机微分方程

假定 $\sigma(x)$，$b(x)$ 足够光滑（即足够多次连续可微）. 我们试图找 Stratonovich 随机微分方程

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)\circ dB_t$$

以 ξ_0 为初始值的解.

假定解的待定形式为（这是关键之处）

$$\xi_t = u(Y_t, B_t),$$

其中 $Y_t = Y_t(\omega)$ 是一个待定的随机过程，且其样本函数 $Y_t(\omega)$ 都是 t 的连续可微函数. 我们要待定合适的函数 $u(y,x)$，及合适的随机过程 Y_t.

为了使 $\xi_t = u(Y_t, B_t)$ 是方程的解，我们用 Ito 公式

$$\mathrm{d}\xi_t = \frac{\partial u}{\partial y}(Y_t,B_t)\frac{\mathrm{d}Y}{\mathrm{d}t}\mathrm{d}t + \left[\frac{\partial u}{\partial x}(Y_t,B_t)\mathrm{d}B_t + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(Y_t,B_t)\mathrm{d}t\right].$$

将它与 Stratonovich 随机微分方程的等价 Ito 随机微分方程

$$\mathrm{d}\xi_t = \left[b(\xi_t) + \frac{1}{2}\sigma'(\xi_t)\sigma(\xi_t)\right]\mathrm{d}t + \sigma(\xi_t)\mathrm{d}B_t$$

作比较,就知道要使 $\xi_t = u(Y_t,B_t)$ 是解,只需满足如下条件:

(A) $\frac{\partial u}{\partial x}(Y_t,B_t) = \sigma(\xi_t) = \sigma(u(Y_t,B_t))$,

(B) $\frac{\partial u}{\partial y}(Y_t,B_t)\frac{\mathrm{d}Y}{\mathrm{d}t} = b(\xi_t) = b(u(Y_t,B_t))$,

(C) $\frac{\partial^2 u}{\partial x^2}(Y_t,B_t) = \sigma'(\xi_t)\sigma(\xi_t) = \sigma'(u(Y_t,B_t))\sigma(u(Y_t,B_t))$.

条件(A)说明 $u(y,x)$ 是非线性常微分方程

$$\begin{cases}\dfrac{\mathrm{d}u}{\mathrm{d}x} = \sigma(u),\\ u(0) = y\end{cases}$$

的解,即

$$u(y,x) = y + \int_0^x \sigma(u(y,a))\mathrm{d}a. \tag{6.9}$$

而且条件(C)也得以满足. 为了由条件(B)确定 Y_t,我们将方程(6.9)对 y 求导数得到

$$\frac{\partial u}{\partial y} = 1 + \int_0^x \sigma'(u(y,a))\frac{\partial u(y,a)}{\partial y}\mathrm{d}a.$$

这说明如果将

$$g(y,x) \xlongequal{\text{def}} \frac{\partial u}{\partial y}$$

视为 x 的函数(y 作为参数),其微商(对 x)就满足方程

$$\begin{cases}g' = \sigma'(u(y,x))g,\\ g(y,0) = 1.\end{cases}$$

这是一个(含参数 y 的)系数为 $\sigma'(u(y,x))$ 的变系数线性常微分方程,故而其解为

$$\frac{\partial u}{\partial y} = g(y,x) = \exp\left[\int_0^x \sigma'(u(y,a))\mathrm{d}a\right].$$

于是条件(B)就等价于

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = b(u(Y_t,B_t))\exp\left[-\int_0^{B_t}\sigma'(u(Y_t,a))\mathrm{d}a\right].$$

对于固定的基本事件 ω 而言,这是 $Y_t(\omega)$ 的一个非线性常微分方程. 此外,在 $t=0$ 时有

$$\xi_0 = u(Y_0,B_0) = u(Y_0,0) = Y_0.$$

也就是说,Y_t 在 $t=0$ 时的初值应该取给定的那个 ξ_0.

更完善的数学处理还必须说明,分别用以确定 $u(y,x)$ 及 $Y_t(\omega)$ 的两个非线性常微分方程的解,以及 Ito 方程的解,都分别在 $-\infty<x<+\infty$ 和 $0\leqslant t<+\infty$ 上有意义. 为此需要验证:$\sigma(u)$,$b(u(y,c))\exp\left[-\int_0^c \sigma'(u(y,a))\mathrm{d}a\right]$($y$ 的函数),$b(u)+\frac{1}{2}\sigma'(u)\sigma(u)$ 都满足局部 Lipshitz 条件和线性增长条件. 这是一个简单的微积分习题,我们略去它的证明.

将上面的推演总结起来,归纳为如下的定理.

定理 6.5　假定 $\sigma(x)$ 有一阶与二阶有界的连续导数,$b(x)$ 有一阶的连续导数. 那么以 ξ_0 为初始值的 Stratonovich 随机微分方程

$$\mathrm{d}\xi_t = b(\xi_t)\mathrm{d}t + \sigma(\xi_t)\circ \mathrm{d}B_t$$

有唯一解,它可以表为

$$\xi_t \overset{\text{def}}{=} u(Y_t, B_t), \tag{6.10}$$

其中 $u(y,x)$ 是非线性常微分方程

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}x} = \sigma(u), \\ u(0) = y \end{cases} \tag{6.11}$$

的解,即满足

$$u(y,x) = y + \int_0^x \sigma(u(y,a))\mathrm{d}a.$$

而随机过程 $Y_t=Y_t(\omega)$ 是将 ω 视为参数的常微分方程

$$\begin{cases} \dfrac{\mathrm{d}Y_t(\omega)}{\mathrm{d}t} = b(u(Y_t,B_t))\exp\left[-\int_0^{B_s} \sigma'(u(Y_t,a))\mathrm{d}a\right], \\ Y_0 = \xi_0 \end{cases} \tag{6.12}$$

的解.

注　定理 6.5 中关于 $b(x)$ 的条件,可以减弱为只需满足局部 Lipshitz 条件,然而,本书中并不追求数学条件的精致.

推论 6.6　假定 $\sigma(x)$ 有一阶与二阶有界的连续导数,$b(x)$ 有一阶的连续导数. 那么以 ξ_0 为初始值的 Ito 随机微分方程

$$\mathrm{d}\xi_t = b(\xi_t)\mathrm{d}t + \sigma(\xi_t)\mathrm{d}B_t$$

$\left(\text{等价于 Stratonovich 随机微分方程 } \mathrm{d}\xi_t=\left[b(\xi_t)-\frac{1}{2}\sigma'(\xi_t)\sigma(\xi_t)\right]\mathrm{d}t+\sigma(\xi_t)\circ\mathrm{d}B_t\right)$ 有唯一解,它可以表示为

$$\xi_t \overset{\text{def}}{=} u(Y_t, B_t),$$

其中 $u(y,x)$ 是非线性常微分方程

$$\begin{cases}\dfrac{\mathrm{d}u}{\mathrm{d}x}=\sigma(u),\\ u(0)=y\end{cases}$$

的解,即满足

$$u(y,x)=y+\int_0^x\sigma(u(y,a))\mathrm{d}a.$$

而随机过程 $Y_t=Y_t(\omega)$ 是将 ω 视为参数的常微分方程

$$\begin{cases}\dfrac{\mathrm{d}Y_t(\omega)}{\mathrm{d}t}=\left[b(u(Y_t,B_t))-\dfrac{1}{2}\sigma'(u(Y_t,B_t))\sigma(u(Y_t,B_t))\right]\exp\left[-\displaystyle\int_0^{B_s}\sigma'(u(Y_t,a))\mathrm{d}a\right],\\ Y_0=\xi_0\end{cases}$$

$$(6.12)'$$

的解.

例 6.13 我们用例 6.7 中的解的表示公式,重新给出例 6.4 中的线性随机微分方程

$$\mathrm{d}\xi_t=(b\xi_t+a)\mathrm{d}t+(\sigma\xi_t+c)\mathrm{d}B_t$$

的解.这里

$$\sigma(x)=\sigma x+c,\quad b(x)=bx+a.$$

方程

$$\frac{\mathrm{d}u}{\mathrm{d}x}=\sigma u+c$$

的通解为

$$u(x)=C\mathrm{e}^{\sigma x}-\frac{c}{\sigma}.$$

因此,以 y 为初始值的解为

$$u(y,x)=y\mathrm{e}^{\sigma x}-\frac{c}{\sigma}(1-\mathrm{e}^{\sigma x}).$$

故而

$$u(Y_t,B_t)=Y_t\exp(\sigma B_t)-\frac{c}{\sigma}[1-\exp(\sigma B_t)].$$

为了确定 Y_t,我们首先注意

$$\begin{aligned}&\left[b(u(Y_t,B_t))-\frac{1}{2}\sigma'(\xi_t)\sigma(\xi_t)\right]\exp\left[-\int_0^{B_t}\sigma'(u(Y_s,a))\mathrm{d}a\right]\\ =&\left[(bu(Y_t,B_t)+a)-\frac{1}{2}\sigma(\sigma u(Y_t,B_t)+c)\right]\exp(-\sigma B_t)\\ =&\left[\left(b-\frac{1}{2}\sigma^2\right)u(Y_t,B_t)+\left(a-\frac{1}{2}\sigma c\right)\right]\exp(-\sigma B_t)\end{aligned}$$

$$=\left(b-\frac{1}{2}\sigma^2\right)\left[Y_t\exp(\sigma B_t)-\frac{c}{\sigma}(1-\exp(\sigma B_t))\right]\exp(-\sigma B_t)$$
$$+\left(a-\frac{1}{2}\sigma c\right)\exp(-\sigma B_t)$$
$$=\left(b-\frac{1}{2}\sigma^2\right)Y_t+\frac{c}{\sigma}\left(b-\frac{1}{2}\sigma^2\right)[1-\exp(-\sigma B_t)]+\left(a-\frac{1}{2}\sigma c\right)\exp(-\sigma B_t).$$

这样,Y_t 满足的线性常微分方程为

$$\frac{dY_t}{dt}=\left(b-\frac{1}{2}\sigma^2\right)Y_t+\frac{c}{\sigma}\left(b-\frac{1}{2}\sigma^2\right)[1-\exp(-\sigma B_t)]+\left(a-\frac{1}{2}\sigma c\right)\exp(-\sigma B_t).$$

我们将它简记为

$$\frac{dY_t}{dt}=A(t)Y_t+f(t).$$

此方程的恰当因子为 $e^{-A(t)}$,通解为

$$Y_t=e^{A(t)}\left(\int_0^t e^{-A(s)}f(s)ds+C\right).$$

初值为 ξ_0 的解就是

$$Y_t=e^{A(t)}\left(\int_0^t e^{-A(s)}f(s)ds+\xi_0\right).$$

于是随机微分方程的解为

$$\xi_t=u(Y_t,B_t)=e^{A(t)}\left(\int_0^t e^{-A(s)}f(s)ds+\xi_0\right)\exp(\sigma B_t)-\frac{c}{\sigma}[1-\exp(\sigma B_t)],$$

其中

$$A(t)=b-\frac{1}{2}\sigma^2,$$
$$f(t)=\frac{c}{\sigma}\left(b-\frac{1}{2}\sigma^2\right)+\left(a-\frac{bc}{\sigma}\right)\exp(-\sigma B_t).$$

*6.3 化简一维随机微分方程的变换方法

6.3.1 可以用变换转化的决定性函数系数的随机微分方程的例子

例 6.14(时变系数的 Black-Scholes 随机微分方程) 对于随机微分方程

$$d\xi_t=b(t)\xi_t dt+\sigma(t)\xi_t dB_t.$$

令 $\eta_t=\ln\xi_t$. 用 Ito 公式得到

$$d\eta_t=\frac{d\xi_t}{\xi_t}-\frac{1}{2}\frac{1}{\xi_t^2}(d\xi_t)^2=\left(b(t)-\frac{\sigma^2(t)}{2}\right)dt+\sigma(t)dB_t.$$

可见

$$\xi_t = \mathrm{e}^{\int_0^t \left[b(s)-\frac{1}{2}\sigma^2(s)\right]\mathrm{d}s+\int_0^t \sigma(s)\mathrm{d}B_s}.$$

例 6.15（随机 Logistic 方程）　Logistic 方程

$$N(t) = rN(t)\left[1-\frac{N(t)}{K}\right]$$

的随机形式是以下的 Ito 方程

$$\mathrm{d}N(t) = N(t)\left[1-\frac{N(t)}{K}\right](r\mathrm{d}t+\alpha\mathrm{d}B(t)).$$

初值 $N(0)=N_0$ 与 Brown 运动$\{B(t):t\geqslant 0\}$独立. 这个方程有两个平凡解：$N(t)\equiv 0$ 或 $N(t)\equiv K$. 这里假定 $0<N(t)<K$. 利用 Ito 公式，有

$$\begin{aligned}
\mathrm{d}\ln\frac{N}{K-N} &= \mathrm{d}\ln N - \mathrm{d}\ln(K-N) \\
&= \left[\frac{\mathrm{d}N}{N}-\frac{1}{2N^2}(\mathrm{d}N)^2\right]-\left[\frac{\mathrm{d}N}{N-K}-\frac{1}{2(N-K)^2}(\mathrm{d}N)^2\right] \\
&= \frac{K}{N(K-N)}\mathrm{d}N-\frac{(K-2N)K}{2N^2(K-N)^2}(\mathrm{d}N)^2 \\
&= r\mathrm{d}t+\alpha\mathrm{d}B-\frac{(K-2N)\alpha^2}{2K}\mathrm{d}t \\
&= \left(r-\frac{1}{2}\alpha^2\right)\mathrm{d}t+\alpha\mathrm{d}B+\frac{N\alpha^2}{K}\mathrm{d}t.
\end{aligned}$$

仿例 6.13 可知

$$\frac{N(t)}{K-N(t)} = C\mathrm{e}^{rt-\frac{1}{2}\alpha^2 t+\alpha B(t)+\frac{\alpha^2}{K}\int_0^t N(s)\mathrm{d}s},$$

其中 $C=\dfrac{N_0}{K-N_0}$. 于是有关系

$$N(t) = \frac{K}{1+\left(\dfrac{K}{N_0-1}\right)\mathrm{e}^{-rt+\frac{1}{2}\alpha^2 t-\alpha B(t)-\frac{\alpha^2}{K}\int_0^t N(s)\mathrm{d}s}}.$$

定义一个函数 $\phi(t)$如下，那么它满足

$$\phi(t) \equiv \frac{K}{1+\left(\dfrac{K}{N_0}-1\right)\mathrm{e}^{-rt+\frac{1}{2}\alpha^2 t-\alpha B(t)}} < N(t) < K.$$

对于

$$x(t) = \ln\frac{N(t)}{K-N(t)}$$

用 Ito 公式得到

$$\mathrm{d}x(t) = \left(r-\frac{1}{2}\alpha^2+\frac{\mathrm{e}^{x(t)}}{1+\mathrm{e}^{x(t)}}\alpha^2\right)\mathrm{d}t+\alpha\mathrm{d}B(t).$$

此方程的系数

$$r-\frac{1}{2}\alpha^2+\frac{e^x}{1+e^x}\alpha^2$$

满足局部 Lipschitz 和线性增长条件，因此具有初值 $x(0)=\ln\dfrac{N_0}{K-N_0}$ 的解是存在且唯一的. 于是

$$N(t)=K\frac{e^{x(t)}}{1+e^{x(t)}}.$$

6.3.2 可以用变换转化的决定性函数系数的随机微分方程的条件

决定性函数为系数的随机微分方程

$$d\xi_t=b(t)dt+\sigma(t)dB_t$$

的解就是随机积分

$$\xi_t=\xi_0+\int_0^t b(t)dt+\int_0^t\sigma(t)dB_t.$$

一个自然的问题是，在什么条件下我们能用变换 $\eta_t=f(t,\xi_t)$ 将随机微分方程

$$d\xi_t=b(t,\xi_t)dt+\sigma(t,\xi_t)dB_t$$

化简为 η_t 的一个确定性系数的随机微分方程

$$d\eta_t=\bar{b}(t)dt+\bar{\sigma}(t)dB_t.$$

为了最终能够解出 ξ_t，我们需要假定 $\eta_t=f(t,\xi_t)$ 有一个反变换

$$\xi_t=g(t,\eta_t).$$

即它们满足

$$\eta_t=f(t,g(t,\eta_t)),\quad \xi_t=g(t,f(t,\xi_t)).$$

我们需要的是待定变换函数的形式 $f(t,x)$ 和系数函数 $\bar{b}(t)$，$\bar{\sigma}(t)$. 为此，对 $\eta_t=f(t,\xi_t)$ 用 Ito 公式，有

$$d\eta_t=f'_t(t,\xi_t)dt+f'_x(t,\xi_t)[b(t,\xi_t)dt+\sigma(t,\xi_t)dB]+\frac{1}{2}f''_{xx}(t,\xi_t)\sigma^2(t,\xi_t)dt.$$

于是我们要求

$$\bar{b}(t)=f'_t(t,x)+f'_x(t,x)b(t,x)+\frac{1}{2}f''_{xx}(t,x)\sigma^2(t,x),$$

$$\bar{\sigma}(t)=f'_x(t,x)\sigma(t,x).$$

由此我们得到下面的引理.

引理 6.7 如果随机微分方程

$$d\xi_t=b(t,\xi_t)dt+\sigma(t,\xi_t)dB_t$$

对于任意 t 固定，存在可逆函数 $f(t,x)$ 使函数

$$f'_t(t,x)+f'_x(t,x)b(t,x)+\frac{1}{2}f''_{xx}(t,x)\sigma^2(t,x)$$

及

$$f'_x(t,x)\sigma(t,x)$$

都不含 x(即只是 t 的函数),将它们分别记为 $\bar{b}(t)$ 及 $\bar{\sigma}(t)$,即

$$\bar{b}(t) = f'_t(t,x) + f'_x(t,x)b(t,x) + \frac{1}{2}f''_{xx}(t,x)\sigma^2(t,x), \tag{6.13}$$

及

$$\bar{\sigma}(t) = f'_x(t,x)\sigma(t,x). \tag{6.14}$$

那么 ξ_t 可以表示为

$$\xi_t = g\left(t, \exp\left[\int_0^t \bar{b}(s)\mathrm{d}s + \int_0^t \bar{\sigma}(s)\mathrm{d}B_s\right]\right),$$

其中 $g(t,x)$ 是 $f(t,x)$ 在 t 固定时的反函数.

下面我们要将存在函数 f 的条件,改进为只用系数描述的条件. 首先由(6.14)式解出

$$f'_x = \frac{\bar{\sigma}(t)}{\sigma(t,x)}, \tag{6.15}$$

求微商得

$$f''_{xx} = -\frac{\bar{\sigma}(t)\sigma_x(t,x)}{\sigma^2(t,x)}.$$

将(6.15)式对 t 求导数得到

$$f''_{xt} = \frac{\bar{\sigma}'(t)\sigma(t,x) - \sigma'_t(t,x)\bar{\sigma}(t)}{\sigma^2(t,x)}.$$

又由于要求(6.13)式不依赖 x,对 x 求微商就得到

$$0 = f''_{tx}(t,x) + \left[f'_x(t,x)b(t,x) + \frac{1}{2}f''_{xx}(t,x)\sigma^2(t,x)\right]'_x.$$

于是

$$0 = \frac{\bar{\sigma}'(t)}{\sigma(t,x)} - \frac{\sigma'_t(t,x)\bar{\sigma}(t)}{\sigma^2(t,x)} + \left[\bar{\sigma}(t)\frac{b(t,x)}{\sigma(t,x)} - \frac{1}{2}\bar{\sigma}(t)\sigma_x(t,x)\right]'_x.$$

也就是

$$\frac{\bar{\sigma}'(t)}{\bar{\sigma}(t)} = \sigma(t,x)\left[\frac{\sigma'_t(t,x)}{\sigma^2(t,x)} - \left[\frac{b(t,x)}{\sigma(t,x)}\right]'_x + \frac{1}{2}\sigma_{xx}(t,x)''\right]. \tag{6.16}$$

因为上式左方的项不依赖 x,故而上式右方对 x 的导数为 0,即方程的系数应该满足条件

$$\left[\sigma(t,x)\left(\frac{\sigma'_t(t,x)}{\sigma^2(t,x)} - \left[\frac{b(t,x)}{\sigma(t,x)}\right]'_x + \frac{1}{2}\sigma_{xx}(t,x)''\right)\right]'_x = 0. \tag{6.17}$$

另一方面,由(6.15)式,存在函数 $C(t)$ 使

$$f(t,x) = \bar{\sigma}(t)\left[\int_1^x \frac{1}{\sigma(t,z)}\mathrm{d}z + C(t)\right]. \tag{6.18}$$

反过来,如果随机微分方程的系数满足(6.17)式,那么,(6.16)式等号的右方的项也

应该与 x 无关. 由(6.16)式积分便得到 $\bar{\sigma}(t)$,例如取

$$\bar{\sigma}(t)=\exp\left\{\int_1^t \sigma(s,x)\left[\frac{\sigma_s'(s,x)}{\sigma^2(s,x)}-\left[\frac{b(s,x)}{\sigma(s,x)}\right]_x'+\frac{1}{2}\sigma_{xx}(s,x)''\right]\mathrm{d}s+M\right\}, \quad (6.19)$$

$$\bar{b}(t)=f_t'(t,x)+f_x'(t,x)b(t,x)+\frac{1}{2}f_{xx}''(t,x)\sigma^2(t,x). \quad (6.20)$$

我们将它归纳为如下的结论.

命题 6.8 若随机微分方程

$$\mathrm{d}\xi_t=b(t,\xi_t)\mathrm{d}t+\sigma(t,\xi_t)\mathrm{d}B_t$$

满足(6.17)式(事实上,由上面的推演可以看到,条件(6.17)也是必要的),那么由(6.20)式,(6.19)式与(6.18)式确定了 $\bar{b}(t)$,$\bar{\sigma}(t)$,$f(t,x)$. 如果对于任意固定的 t,$f(t,x)$的反函数 $g(t,x)$存在. 那么有

$$\xi_t=g\left(t,\int_0^t \bar{b}(s)\mathrm{d}s+\int_0^t \bar{\sigma}(s)\mathrm{d}B_s\right).$$

例 6.16 对于例 6.14,有

$$b(t,x)=b(t)x,\quad \sigma(t,x)=\sigma(t)x.$$

$$\bar{\sigma}(t)=\exp\left(\int_1^t \frac{\sigma'(s)}{\sigma(s)}\mathrm{d}s+M\right)\overset{\text{def}}{=}\sigma(t),$$

$$f(t,x)=\frac{\bar{\sigma}(t)}{\sigma(t)}\left[\int_1^x \frac{1}{z}\mathrm{d}z+C(t)\right]\overset{\text{def}}{=}\frac{\bar{\sigma}(t)}{\sigma(t)}\ln x=\ln x,$$

那么

$$\bar{b}(t)=b(t)-\frac{1}{2}\sigma^2(t).$$

所以

$$\xi_t=\xi_0\exp\left\{\int_0^t\left[b(s)-\frac{1}{2}\sigma^2(s)\right]\mathrm{d}s+\int_0^t\sigma(s)\mathrm{d}B_s\right\}.$$

下面我们讨论系数不含 t 的情形. 如果对于随机微分方程

$$\mathrm{d}\xi_t=b(\xi_t)\mathrm{d}t+\sigma(\xi_t)\mathrm{d}B_t,$$

存在可逆函数 $f(x)$使函数

$$f'(x)b(x)+\frac{1}{2}f''(x)\sigma^2(x) \quad (6.13)'$$

及

$$f'(x)\sigma(x) \quad (6.14)'$$

都不含 x,分别记为$\bar{b}$,$\bar{\sigma}$,那么取

$$f(x)=\bar{\sigma}\left(\int_1^x \frac{1}{\sigma(z)}\mathrm{d}z+C\right),$$

$$\bar{b}=\bar{\sigma}\left(\frac{b(x)}{\sigma(x)}-\frac{1}{2}\sigma'(x)\right).$$

故而

$$\frac{1}{2}\sigma(x)' - \frac{b(x)}{\sigma(x)} = -\frac{\bar{b}}{\bar{\sigma}} \stackrel{\text{def}}{=} \gamma.$$

即$\bar{b} = \gamma\bar{\sigma}$，从而有

$$\mathrm{d}f(\xi_t) = \bar{\sigma}(\gamma\mathrm{d}t + \mathrm{d}B_t).$$

特别地，可以取 $\bar{\sigma}=1$. 所以我们得到了如下的推论.

推论 6.9（系数不含 t 的情形） 若随机微分方程

$$\mathrm{d}\xi_t = b(\xi_t)\mathrm{d}t + \sigma(\xi_t)\mathrm{d}B_t$$

满足

$$\frac{1}{2}\sigma(x)' - \frac{b(x)}{\sigma(x)} \equiv \text{常数 } \gamma,$$

令

$$f(x) = \int_1^x \frac{1}{\sigma(z)}\mathrm{d}z + C,$$

$$\bar{b} = \frac{b(x)}{\sigma(x)} - \frac{1}{2}\sigma'(x).$$

那么我们可以用变换 $\eta_t = f(\xi_t)$将 ξ_t 随机微分方程转化为常系数随机微分方程

$$\mathrm{d}\eta_t = \gamma\mathrm{d}t + \mathrm{d}B_t.$$

6.3.3 求解 Ito 方程的另一种途径——将方程中 Brown 运动的系数化为 1

考虑随机微分方程

$$\mathrm{d}\xi_t = b(\xi_t)\mathrm{d}t + \sigma(\xi_t)\mathrm{d}B_t,$$

它可以化为 Stratonovich 随机微分方程，然后在原则上可以用定理 6.5 求解. 本节我们考虑另一种方法，即选取变换

$$\xi_t = u(\eta_t) \quad (\xi_0 = u(\eta_0)),$$

使原方程化为

$$\mathrm{d}\eta_t = g(\xi_t)\mathrm{d}t + \mathrm{d}B_t,$$

而对后者再用定理 6.5 的方法求解有时可能比原方程用定理 6.5 容易求解. 为了选取函数 $u(x)$，我们用 Ito 公式

$$\begin{aligned}\mathrm{d}\xi_t &= u'(\eta_t)\mathrm{d}\eta_t + \frac{1}{2}u''(\eta_t)\mathrm{d}t \\ &= \left[u'(\eta_t)g(\eta_t) + \frac{1}{2}u''(\eta_t)\right]\mathrm{d}t + u'(\eta_t)\mathrm{d}B_t,\end{aligned}$$

故而应该有

$$u'(\eta_t)g(\eta_t) + \frac{1}{2}u''(\eta_t) = b(\xi_t) = b(u(\eta_t)),$$

$$u'(\eta_t) = \sigma(\xi_t) = \sigma(u(\eta_t)).$$

为此只需取 $u(x)$ 满足非线性常微分方程

$$\begin{cases} u'(y) = \sigma(u(y)), \\ u(\eta_0) = \xi_0. \end{cases}$$

并且取

$$g(y) = \frac{b(u(y)) - \frac{1}{2}u''(y)}{\sigma(u(y))}$$

即可.

6.4 随机微分方程解的矩与对参数的依赖

下列两个定理在应用中相当重要,其证明在理论较多一些的教科书上可以找到,我们略去其证明.

定理 6.10 如果 d 维随机微分方程

$$\mathrm{d}\xi_t = b(t,\boldsymbol{\xi}_t)\mathrm{d}t + \boldsymbol{\Sigma}(t,\boldsymbol{\xi}_t)\mathrm{d}B_t$$

的系数 $\boldsymbol{b}(t,\boldsymbol{x}),\boldsymbol{\Sigma}(t,\boldsymbol{x})=(\sigma_{ik}(t,\boldsymbol{x}))_{i\leqslant d,k\leqslant m}$满足相应于定理 6.1 的条件,而且初值为$\boldsymbol{\xi}_0$ 具有 $2p$ 阶矩

$$E|\boldsymbol{\xi}_0|^{2p} < +\infty \qquad (p>0).$$

那么存在常数 α_T,β_T 使其解$\boldsymbol{\xi}_t$ 在 $t\leqslant T$ 满足

$$E|\boldsymbol{\xi}_t|^{2p} \leqslant \alpha_T(1+E|\boldsymbol{\xi}_0|^{2p})\mathrm{e}^{\beta_T t},$$
$$E|\boldsymbol{\xi}_t-\boldsymbol{\xi}_0|^{2p} \leqslant \alpha_T t^p(1+E|\boldsymbol{\xi}_0|^{2p})\mathrm{e}^{\beta_T t}.$$

定理 6.11(系数扰动下的稳定性) 如果初值为$\boldsymbol{\xi}_0^{(n)}$,d 维随机微分方程序列

$$\mathrm{d}\boldsymbol{\xi}_t = \boldsymbol{b}_n(\boldsymbol{\xi}_t)\mathrm{d}t + \boldsymbol{\Sigma}_n(\boldsymbol{\xi}_t)\mathrm{d}\boldsymbol{B}_t$$

的系数满足相应于定理 6.1 的条件,而且还满足

$$\lim_{n\to+\infty} E|\boldsymbol{\xi}_0^{(n)}-\boldsymbol{\xi}_0|^2 = 0,$$

对于任意 $K>0$

$$\lim_{n\to+\infty}\max_{t\leqslant T,|\boldsymbol{x}|\leqslant K}\{|\boldsymbol{b}_n(t,\boldsymbol{x})-\boldsymbol{b}(t,\boldsymbol{x})|+|\boldsymbol{\Sigma}_n(t,\boldsymbol{x})-\boldsymbol{\Sigma}(t,\boldsymbol{x})|\} = 0,$$

那么

$$\lim_{n\to+\infty} E(\max_{t\leqslant T}|\boldsymbol{\xi}_t^{(n)}-\boldsymbol{\xi}_t|)^2 = 0.$$

其中$\boldsymbol{\xi}_t$ 是随机微分方程

$$\mathrm{d}\boldsymbol{\xi}_t = \boldsymbol{b}(t,\boldsymbol{\xi}_t)\mathrm{d}t + \boldsymbol{\Sigma}(t,\boldsymbol{\xi}_t)\mathrm{d}\boldsymbol{B}_t$$

以$\boldsymbol{\xi}_0$为初值的解.

这个定理的典型应用之一,是常微分方程在小扰动下解的稳定性:

$$\begin{cases} \mathrm{d}\boldsymbol{\xi}_t = \boldsymbol{b}(\boldsymbol{\xi}_t)\mathrm{d}t + \varepsilon \mathrm{d}\boldsymbol{B}_t \\ \boldsymbol{\xi}_t\Big|_{t=0} = \boldsymbol{x}_0 \end{cases}$$

的解$\boldsymbol{\xi}_t^{(\varepsilon)}$对于任意$T>0$,在$t\leqslant T$有

$$\lim_{\varepsilon\to 0} E(\max_{t\leqslant T}|\boldsymbol{\xi}_t^{(\varepsilon)} - \boldsymbol{x}_t|)^2 = 0,$$

其中$\boldsymbol{x}_t$是常微分方程

$$\begin{cases} \mathrm{d}\boldsymbol{x}_t = \boldsymbol{b}(\boldsymbol{x}_t)\mathrm{d}t \\ \boldsymbol{x}_{t|t=0} = \boldsymbol{x}_0 \end{cases}$$

的解.

6.5　Kalman-Bucy 滤波

6.5.1　Gauss 过程的投影——线性滤波

定义 6.2　设$(\xi_t: t\in I)\cup\{\eta\}$是期望函数为 0 的 Gauss 系.随机变量$\eta$在$\{\xi_t: t\in I\}$的非线性均方信息空间上的投影(我们将非线性均方信息空间,记为$L^2(\bar{\Phi}(\xi))$,它就是在$\{\xi_t: t\in I\}$的非线性信息空间$\bar{\Phi}(\xi)$中具有有限方差的随机变量全体)记为$\check{\eta}$.它可以由如下的两个条件唯一地确定:

$$\check{\eta}\in L^2(\bar{\Phi}(\xi)),$$

$$\text{对于任意 }\zeta\in L^2(\bar{\Phi}(\xi)),\text{都有 }E[(\eta-\hat{\eta})\bar{\zeta}]=0(\textbf{非线性投影公式}).\qquad(6.21)$$

即在内积含义下,$\eta-\check{\eta}\perp L^2(\bar{\Phi}(\xi))$.$\check{\eta}$称为$\eta$关于已知信息$\{\xi_t: t\in I\}$的**非线性滤波**.

随机变量η在$\{\xi_t: t\in I\}$的线性均方信息空间$\bar{L}(\xi)$上的投影,记为$\hat{\eta}$.它可以由如下的两个条件唯一地确定

$$\hat{\eta}\in\bar{L}(\xi),$$

$$\text{对于任意 }\zeta\in\bar{L}(\xi),\text{都有 }E[(\eta-\hat{\eta})\bar{\zeta}]=0(\textbf{线性投影公式}),\qquad(6.22)$$

即在内积含义下,$\eta-\hat{\eta}\perp\bar{L}(\xi)$.$\hat{\eta}$称为$\eta$关于已知信息$\{\xi_t: t\in I\}$的**线性滤波**,也记为$\mathrm{Proj}_{\bar{L}(\xi)}\eta$.

定理 6.12　设$\{\xi_t: t\in I\}\cup\{\eta\}$是期望为 0 的 Gauss 系,那么

$$\check{\eta}=\hat{\eta},$$

即对 Gauss 系而言,非线性滤波与线性滤波是一样的.

证明　首先注意$\hat{\eta}\in\bar{L}(\xi)\subset L^2(\bar{\Phi}(\xi))$.其次,由于$\eta-\hat{\eta}\perp\bar{L}(\xi)$,有$E[(\eta-\hat{\eta})\xi_t]=0$

$(\forall t\in I)$. 由命题 1.9 得到，$\eta-\hat{\eta}$ 与 $\{\xi_t: t\in I\}$ 独立，因而也与 $L^2(\bar{\Phi}(\xi))$ 独立. 于是对于任意对 $\zeta\in L^2(\bar{\Phi}(\xi))$ 有 $E[(\eta-\hat{\eta})\zeta]=0$. 这正说明了 $\eta-\hat{\eta}\perp L^2(\bar{\Phi}(\xi))$. 从而 $\check{\eta}=\hat{\eta}$.

6.5.2 Kalman-Bucy 滤波模型与滤波过程（用新息过程表示，滤波的均方误差），滤波的随机微分方程

背景 在观测信号中混有噪声干扰时，就要利用信号本身的规律去过滤掉噪声，这就是滤波问题. 在很多情形下信号 ξ_t 无法直接测量，而能实测到的数据是由信号在空间传播的方式与随机噪声 W_t 决定的. 假定信号是通过多个通道接收的，而且到达各接收点时间与强度并不一致. 这样，汇总而接收到的测量值，并不是信号 ξ_t，而是信号的现在值与它过去多个时刻的值的不同权重的滑动，再加上随机噪声：

$$\sum G_{t_j}\xi_{t-t_j}+W_t.$$

我们将它记为 η_t，即

$$\eta_t=\sum G_{t_j}\xi_{t-t_j}+W_t,$$

它称为量测方程. 连续信息的 Kalman-Bucy 滤波模型是多个时刻的信息滤波的积分近似，需要用随机微分方程的概念，并考虑如下的量测方程

$$\mathrm{d}\eta_t=G(t)\mathrm{d}\xi_t+D(t)\mathrm{d}W_t.$$

Kalman-Bucy 滤波模型

设 $\{B_t: t\geqslant 0\}$，$\{W_t: t\geqslant 0\}$ 是两个彼此独立的 Brown 运动，而状态随机过程 ξ_t 满足如(6.5)式的线性随机微分方程

$$\mathrm{d}\xi_t=F(t)\xi_t\mathrm{d}t+C(t)\mathrm{d}B_t$$

称为**状态方程**. 虽然(6.5)式给出了状态过程 ξ_t 的表达式，但是在有干扰环境的实际问题中，状态过程 ξ_t 无法直接测量得到(从模型角度解释，就是在(6.5)式中的参变 Brown 运动不知道是哪一个). 而能测量到的是与状态过程 ξ_t 有关的另一个 Ito 过程

$$\mathrm{d}\eta_t=G(t)\xi_t\mathrm{d}t+D(t)\mathrm{d}W_t \tag{6.23}$$

称为**量测方程**，其中 $F(t),C(t),G(t),D(t)(D(t)>0)$ 都是非随机的数值函数. 这时状态过程与量测过程都是 Ito 过程. 这个由一对状态过程与量测过程组成的模型，称为连续时间的 **Kalman-Bucy 滤波模型**.

(1) 新息过程

因为对于任意固定的 t，随机过程 $\{\eta_t,(\xi_s,s\leqslant t)\}$ 是 Gauss 过程，所以关于它的非线性滤波与线性滤波是一样的. 从而我们在下文中，只需考虑线性滤波.

定义 6.3 我们将

$$\hat{\xi}_t\overset{\text{def}}{=\!=}\mathrm{Proj}\bar{L}(\eta_s,s\leqslant t)\xi_t=E(\xi_t\mid\eta_s,s\leqslant t) \tag{6.24}$$

称为**状态的滤波过程**,或者**滤波器**. 并将

$$\tilde{\xi}_t \stackrel{\text{def}}{=} \xi_t - \hat{\xi}_t \tag{6.25}$$

称为**滤波误差过程**. 而将

$$P(t) \stackrel{\text{def}}{=} E\tilde{\xi}_t^2 \tag{6.26}$$

称为**滤波的均方误差函数**. 进一步将 Ito 过程

$$n_t \stackrel{\text{def}}{=} \eta_t - \int_0^t G(s)\hat{\xi}_s \mathrm{d}s = \int_0^t G(s)(\xi_s - \hat{\xi}_s)\mathrm{d}s + \int_0^t D(s)\mathrm{d}W_s \tag{6.27}$$

称为滤波问题的**新息过程**. 它代表量测过程随着时间增长提供的新的补充信息.

如果我们将随机系$\{n_s: s\leqslant t\}$的线性均方信息空间(回忆起 1.3 节中记号的含义,它是对于极限运算封闭的),记为$\bar{L}\{n_s: s\leqslant t\}$,那么,我们有下面的引理.

引理 6.13 随机系$\{\eta_s: s\leqslant t\}$的线性均方信息空间正是数值函数的随机积分全体:

$$\bar{L}\{\eta_s: s\leqslant t\} = \left\{\int_0^t f(s)\mathrm{d}\eta_s: \int_0^t f^2(s)\mathrm{d}s < +\infty\right\}. \tag{6.28}$$

特别地,对于任意 $t>0$,存在函数 $h_t(s)(s\leqslant t)$使

$$\hat{\xi}_t = \int_0^t h_t(s)\mathrm{d}\eta_s. \tag{6.29}$$

证明 对于任意 $\int_0^t f^2(s)\mathrm{d}s < +\infty$,因为随机积分$\int_0^t f(s)\mathrm{d}\eta_s$ 是随机和的极限,而随机和属于 $\bar{L}\{\eta_s: s\leqslant t\}$,所以$\int_0^t f(s)\mathrm{d}\eta_s \in \bar{L}\{\eta_s: s\leqslant t\}$. 故而

(6.28)式等号右边$\subset$(6.28)式等号左边.

另一方面,对于任意 $s\leqslant u$,有

$$\eta_t = \int_0^t I_{[0,u]}(s)\mathrm{d}\eta_s \in (6.28)\text{ 式等号右边.}$$

但是(6.28)式等号右边对于线性运算及极限封闭,因此

(6.28)式等号左边$\subset$(6.28)式等号右边.

引理 6.14 量测过程$\{\eta_s: s\leqslant t\}$也可以通过新息过程$\{n_s: s\leqslant t\}$表示,即对于任意 $t>0$,存在数值函数($f_t(s)$,又因为 t 是固定的,在不至于混淆的情形下,我们可以将它简写为)$f(s)$,使

$$\eta_t = \int_0^t f(s)\mathrm{d}n_s.$$

证明 我们先分析这个表达式成立的条件

$$\begin{aligned}\int_0^t f(s)\mathrm{d}n_s &= \int_0^t f(s)\mathrm{d}\eta_s - \int_0^t f(s)G(s)\hat{\xi}_s\mathrm{d}s \\ &= \int_0^t f(s)\mathrm{d}\eta_s - \int_0^t f(s)G(s)\int_0^s h_s(u)\mathrm{d}\eta_u\mathrm{d}s\end{aligned}$$

$$= \int_0^t f(s)\mathrm{d}\eta_s - \int_0^t \left[\int_u^t f(s)h(s,u)\mathrm{d}s\right]\mathrm{d}\eta_u \quad (\text{其中 } h(t,s) = G(t)h_t(s), s \leqslant t)$$

$$= \int_0^t f(s)\mathrm{d}\eta_s - \int_0^t \left[\int_s^t f(v)h(v,s)\mathrm{d}v\right]\mathrm{d}\eta_s \quad (\text{哑变量换一个记号}).$$

所以,能否有表示式 $\eta_t = \int_0^t f(s)\mathrm{d}n_s$ 的关键在于,是否存在 $f(s)$ 使

$$\eta_t = \int_0^t f(s)\mathrm{d}\eta_s - \int_0^t \left[\int_s^t h(v,s)\mathrm{d}v\right]\mathrm{d}\eta_s.$$

即用以待定函数 $f(s)(s\leqslant t)$的方程

$$I_{[0,t]}(s) = f(s) - \int_s^t f(v)h(v,s)\mathrm{d}v$$

是否存在一个解. 事实上,这个方程是 Volterra 积分方程,不仅存在唯一解,而且其解有级数表示. 这样我们就证明了:存在 $f(s)(s\leqslant t)$,使 $\eta_t = \int_0^t f(s)\mathrm{d}n_s$.

命题 6.15 新息过程$\{n_t: t\geqslant 0\}$是 Gauss 过程. 而且满足:

(1) $\{n_t: t\geqslant 0\}$是正交增量过程,即对于任意 $t_1<t_2\leqslant t_3<t_4$,恒有

$$E[(n_{t_2} - n_{t_1})(n_{t_4} - n_{t_3})] = 0,$$

从而也是独立增量过程.

(2)

$$En_t = 0, \quad En_t^2 = \int_0^t D(s)^2\mathrm{d}s.$$

(3) 新息过程$\{n_t: t\geqslant 0\}$与量测过程$\{\eta_t: t\geqslant 0\}$有相同的历史信息,即随机系$\{n_s: s\leqslant t\}$和$\{\eta_s: s\leqslant t\}$的线性均方信息空间 $\bar{L}\{n_s: s\leqslant t\}$和 $\bar{L}\{\eta_s: s\leqslant t\}$是相同的.

证明 新息过程

$$\left\{n_t = \eta_t - \int_0^t G(s)\hat{\xi}_s\mathrm{d}s, t\geqslant 0\right\}$$

可以写成和的极限,故由定理 1.10 推得它是 Gauss 过程.

(1) 对于任意 $s<t$ 及任意$\{\eta_u: u\leqslant s\}$可知的有界随机变量 ζ,由射影的定义可得

$$E[(n_t - n_s)\zeta] = E\left(\left[\int_s^t G(u)(\xi_u - \hat{\xi}_u)\mathrm{d}u + \int_s^t D(u)\mathrm{d}W_u\right]\zeta\right).$$

$$= \int_s^t G(u)E[(\xi_u - \hat{\xi}_u)\zeta]\mathrm{d}u = 0.$$

用一点数学推导就可以说明,对于线性均方信息空间 $\bar{L}\{n_s: s\leqslant t\}$中的 ζ,仍然有

$$E[(n_t - n_s)\zeta] = 0.$$

取 $s=t_3, t=t_4, \zeta=n_{t_2}-n_{t_1}$,就得到信息过程的正交增量性.

(2) 首先,$En_t = \int_0^t G(s)E(\xi_s - \hat{\xi}_s)\mathrm{d}s = 0$. 又因为信息过程$\{n_t: t\geqslant 0\}$也是 Ito 过程

$$\mathrm{d}n_t = \mathrm{d}\eta_t - G(t)\hat{\xi}_t\mathrm{d}t = G(t)(\xi_t - \hat{\xi}_t)\mathrm{d}t + D(t)\mathrm{d}W_t,$$
$$(\mathrm{d}n_t)^2 = (\mathrm{d}\eta_t)^2 = [D(t)]^2\mathrm{d}t.$$

用 Ito 公式得到

$$\mathrm{d}n_t^2 = 2n_t\mathrm{d}n_t + (\mathrm{d}n_t)^2 = 2G(t)(\xi_t - \hat{\xi}_t)n_t\mathrm{d}t + [D(t)]^2\mathrm{d}t,$$

即

$$n_t^2 = 2\int_0^t G(t)(\xi_s - \hat{\xi}_s)n_s\mathrm{d}s + \int_0^t [D(s)]^2\mathrm{d}s.$$

故而

$$En_t^2 = 2\int_0^t G(t)E[(\xi_s - \hat{\xi}_s)n_s]\mathrm{d}s + \int_0^t [D(s)]^2\mathrm{d}s = \int_0^t [D(s)]^2\mathrm{d}s.$$

(3) 由定义新息过程是(η_t)可知的. 故线性均方信息空间 $\bar{L}\{n_s: s\leqslant t\}\subset\bar{L}\{\eta_s: s\leqslant t\}$. 而反向的包含关系只是引理 6.7 的一个直接推论.

命题 6.16 定义规范新息过程为

$$V_t \overset{\text{def}}{=} \int_0^t \frac{1}{D(s)}\mathrm{d}n_s = \int_0^t \frac{G(s)}{D(s)}(\xi_s - \hat{\xi}_s)\mathrm{d}s + W_t. \tag{6.30}$$

那么, $\{V_t, t\geqslant 0\}$是一个 Brown 运动.

证明 因为新息过程$\{n_t, t\geqslant 0\}$是正交增量的 Gauss 过程, 所以$\{V_t, t\geqslant 0\}$也是正交增量的 Gauss 过程, 从而也是独立增量过程. 而且有 $EV_t=0$. 另一方面, 由(6.30)式推出

$$EV_t^2 = \int_0^t \frac{1}{D(s)^2}En_s^2\mathrm{d}s = \int_0^t \frac{1}{D(s)^2}D(s)^2\mathrm{d}s = t.$$

因此$\{V_t, t\geqslant 0\}$是 Brown 运动.

(2) 用均方误差函数与规范新息过程表达滤波过程

引理 6.17 对于滤波过程 $\hat{\xi}_t$ 成立如下诸关系:

(1) 商高定理

$$E(\xi_t - \hat{\xi}_t)^2 = E\xi_t^2 - E\hat{\xi}_t^2. \tag{6.31}$$

(2)

$$\hat{\xi}_t = \int_0^t g_t(s)\mathrm{d}V_s, \tag{6.32}$$

$$E(\xi_t V_s) = \int_0^t g_t(s)\mathrm{d}s,$$

其中

$$g_t(s) = \frac{G(s)}{D(s)}P(s)\mathrm{e}^{\int_s^t F(u)\mathrm{d}u}, \tag{6.33}$$

而 $P(t)$是滤波的均方误差函数.

证明 由(6.29)式

$$\hat{\xi}_t = \int_0^t h_t(s)\mathrm{d}n_s = \int_0^t h_t(s)D(s)\mathrm{d}V_s.$$

令

$$g_t(s) \overset{\text{def}}{=} h_t(s)D(s),$$

便得(6.32)式. 再则,因为滤波过程 $\hat{\xi}_t$ 是线性均方信息空间 $\bar{L}\{\eta_s: s\leqslant t\}$ 中的元素,以及

$$\bar{L}\{V_s: s\leqslant t\} = \bar{L}\{n_s: s\leqslant t\} = \bar{L}\{\eta_s: s\leqslant t\},$$

从而 $\hat{\xi}_t$ 是线性均方信息空间 $\bar{L}\{V_s: s\leqslant t\}$ 中的元素. 根据投影的性质,对于 $s\leqslant t$ 有

$$E[(\xi_t - \hat{\xi}_t)V_s] = 0, \quad E[(\xi_t - \hat{\xi}_t)\hat{\xi}_s] = 0.$$

故而得到

$$E(\xi_t V_s) = E(\hat{\xi}_t V_s) = \int_0^s g_t(u)\mathrm{d}u$$

及

$$E(\xi_t - \hat{\xi}_t)^2 = E[\xi_t(\xi_t - \hat{\xi}_t)] = E\xi_t^2 - E\hat{\xi}_t^2.$$

于是

$$g_t(s) = \frac{\mathrm{d}}{\mathrm{d}s}E(\xi_t V_s).$$

利用状态过程 $\{\xi_t, t\geqslant 0\}$ 与 Brown 运动 $\{W_t, t\geqslant 0\}$ 相互独立及(6.27)式,推出

$$\begin{aligned} E(\xi_t V_s) &= E\left[\xi_t\left(\int_0^t \frac{G(s)}{D(s)}(\xi_s - \hat{\xi}_s)\mathrm{d}s + W_t\right)\right] \\ &= \int_0^t \frac{G(s)}{D(s)}E[\xi_t(\xi_s - \hat{\xi}_s)]\mathrm{d}s. \end{aligned}$$

另一方面,由(6.5)式及(6.6)式有

$$\xi_t = \xi_s \mathrm{e}^{\int_s^t F(u)\mathrm{d}u} + \int_s^t C(u)\mathrm{e}^{\int_u^t F(v)\mathrm{d}v}\mathrm{d}B_u.$$

可知 ξ_t 是独立增量过程,这样就得到

$$E(\xi_t V_s) = \int_0^t \frac{G(s)}{D(s)}E[\xi_s \mathrm{e}^{\int_s^t F(u)\mathrm{d}u}(\xi_s - \hat{\xi}_s)]\mathrm{d}s = \int_0^t \frac{G(s)}{D(s)}P(s)\mathrm{e}^{\int_s^t F(u)\mathrm{d}u}\mathrm{d}s.$$

从而导出了函数 $g_t(s)$ 的表达式为

$$g_t(s) = \frac{G(s)}{D(s)}P(s)\mathrm{e}^{\int_s^t F(u)\mathrm{d}u}.$$

(3) 滤波均方误差函数与其 Riccati 方程

定理 6.18 滤波的均方误差函数 $P(t)$是 Riccati 方程

$$\begin{cases} \dfrac{\mathrm{d}P(t)}{\mathrm{d}t} = 2F(t)P(t) - \dfrac{G^2(t)}{D^2(t)}P^2(t) + C^2(t), \\ P(0) = \mathrm{var}(\xi_0) \end{cases} \tag{6.34}$$

的解.

证明　对于滤波的均方误差函数,利用(6.6)式与(6.31)～(6.33)诸式推出

$$P(t)=E(\xi_t-\hat{\xi}_t)^2=E\xi_t^2-E\hat{\xi}_t^2=E\xi_0^2\mathrm{e}^{2\int_0^t F(u)\mathrm{d}u}$$
$$+\int_0^t C^2(u)\mathrm{e}^{2\int_u^t F(v)\mathrm{d}v}\mathrm{d}u-\int_0^t\left(\frac{G(u)}{D(u)}\right)^2 P^2(u)\mathrm{e}^{2\int_u^t F(v)\mathrm{d}v}\mathrm{d}u.$$

求导数,再由上式得到

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t}=2F(t)E\xi_0^2\mathrm{e}^{2\int_0^t F(u)\mathrm{d}u}+C^2(t)+2F(t)\int_0^t C^2(u)\mathrm{e}^{2\int_u^t F(v)\mathrm{d}v}\mathrm{d}u$$
$$-\left(\frac{G(t)}{D(t)}\right)^2 P^2(t)-2F(t)\int_0^t\left(\frac{G(u)}{D(u)}\right)^2 P^2(u)\mathrm{e}^{2\int_u^t F(v)\mathrm{d}v}\mathrm{d}u.$$

最后化简为

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t}=2F(t)P(t)+C^2(t)-\frac{G^2(t)}{D^2(t)}P^2(t).$$

这是一个 Riccatti 方程,其初值条件为($\hat{\xi}_0\stackrel{\text{def}}{=\!=}E\xi_0$)

$$P(0)=\mathrm{var}(\xi_0).$$

(4) 滤波过程满足的随机微分方程

对于(6.32)式中的滤波过程 $\hat{\xi}_t$,改记

$$g(t,s)\stackrel{\text{def}}{=\!=}g_t(s).$$

用例 5.9 及(6.32),(6.33)式得到

$$\begin{aligned}\mathrm{d}\hat{\xi}_t&=g(t,t)\mathrm{d}V_t+\int_0^t\frac{\partial g}{\partial t}(t,s)\mathrm{d}V_s\mathrm{d}t\\&=\frac{G(t)}{D(t)}P(t)\mathrm{d}V_t+F(t)\int_0^t g(t,s)\mathrm{d}V_s\mathrm{d}t\\&=\frac{G(t)}{D(t)}P(t)\left[\frac{G(t)}{D(t)}(\xi_t-\hat{\xi}_t)\mathrm{d}t+\mathrm{d}W_t\right]+F(t)\hat{\xi}_t\mathrm{d}t\\&=\frac{G(t)}{D^2(t)}P(t)[\mathrm{d}\eta_t-G(t)\hat{\xi}_t\mathrm{d}t]+F(t)\hat{\xi}_t\mathrm{d}t.\end{aligned}$$

所以我们证明了以下定理.

定理 6.19　滤波过程是随机微分方程

$$\mathrm{d}\hat{\xi}_t=\left[F(t)-\frac{G^2(t)}{D^2(t)}P(t)\right]\hat{\xi}_t\mathrm{d}t+\frac{G(t)}{D^2(t)}P(t)\mathrm{d}\eta_t\qquad(6.35)$$

的解,这是一个非随机时变系数的线性随机微分方程,称为滤波过程的 **Kalman-Bucy 随机微分方程.**

方程(6.35)也可以写为

$$\mathrm{d}\hat{\xi}_t=F(t)\hat{\xi}_t\mathrm{d}t+K(t)[\mathrm{d}\eta_t-G(t)\hat{\xi}_t\mathrm{d}t],\qquad(6.35)'$$

或

$$d\hat{\xi}_t = F(t)\hat{\xi}_t dt + K(t)dn_t, \tag{6.35}''$$

其中

$$K(t) \overset{\text{def}}{=\!=} \frac{G(t)}{D^2(t)}P(t) \tag{6.36}$$

称为 Kalman-Bucy 滤波的**增益系数**，于是，Kalman-Bucy 滤波方程有明显的直观含义：即滤波估计 $\hat{\xi}_t$ 的更新量 $d\hat{\xi}_t$ 由两部分构成：按照状态过程作更新的部分 $F(t)\hat{\xi}_t dt$，和新息过程 n_t 提供作更新的部分 $K(t)dn_t$，而增益系数正是新息过程给滤波更新部分的瞬时强度.

在作实际的滤波时，首先是建模，最主要的是确定 $F(t), C(t), G(t), D(t), E\xi_0$, $\mathrm{var}(\xi_0)$. 其次是求解 Riccatti 方程. 只要得到了 Riccatti 方程的解 $P(t)$，那么，就可以通过解 Kalman-Bucy 滤波方程得到滤波过程的解析表达式

$$\begin{aligned}\hat{\xi}_t = &(E\xi_0)\exp\left(\int_0^t \left[F(u) - \frac{G^2(u)}{D^2(u)}P(u)\right]du\right) \\ &+ \int_0^t \frac{G(s)}{D^2(s)}P(s)\exp\left(\int_s^t \left[F(u) - \frac{G^2(u)}{D^2(u)}P(u)\right]du\right)d\eta_s.\end{aligned} \tag{6.37}$$

从而也可以作数值的近似计算.

例 6.17(常系数情形)　系数不依赖 t 时，决定滤波均方误差函数 $P(t)$ 的 Riccatti 方程为

$$\begin{cases}\dfrac{dP(t)}{dt} = 2FP(t) - \dfrac{G^2}{D^2}P^2(t) + C^2, \\ P(0) = \mathrm{var}(\xi_0).\end{cases}$$

常系数 Riccatti 方程的解有公式表示，在常微分方程的一些教程上可以找到. 这个解的表示式为(读者可以直接验证它是解)

$$P(t) = \frac{\alpha - \kappa\beta\exp\left((\beta-\alpha)\dfrac{G^2}{D^2}t\right)}{1 - \kappa\exp\left((\beta-\alpha)\dfrac{G^2}{D^2}t\right)} \approx \beta(\text{与 } t \text{ 大时}), \tag{6.38}$$

其中

$$\kappa = \frac{P(0)-\alpha}{P(0)-\beta},$$

$$\alpha = \frac{1}{G^2}(FD^2 - D\sqrt{F^2D^2 + G^2C^2}),$$

$$\beta = \frac{1}{G^2}(FD^2 + D\sqrt{F^2D^2 + G^2C^2}).$$

于是得到滤波过程的表达式

$$\hat{\xi}_t = (E\xi_0)\exp\left(\int_0^t \left[F - \frac{G^2}{D^2}P(u)\right]du\right) + \frac{G}{D^2}\int_0^t P(s)\exp\left(\int_s^t \left[F - \frac{G^2}{D^2}P(u)\right]du\right)d\eta_s$$

$$\approx (E\xi_0)\exp\left(\left[F - \frac{G^2}{D^2}\beta\right]t\right) + \frac{G}{D^2}\beta\int_0^t \exp\left(\left[F - \frac{G^2}{D^2}\beta\right](t-s)d\eta_s\right.(\text{与 } t \text{ 大时}). \quad (6.39)$$

例 6.18（未知参数的估计）　假设需要用观测过程

$$d\eta_t = \theta G(t)dt + D(t)dW_t$$

去估计它包含的未知参数 θ，那么，状态方程可以取

$$d\theta = 0.$$

于是

$$F(t) = C(t) = 0, \quad \xi_t \equiv \theta.$$

决定滤波均方误差函数 $P(t)$ 的 Riccati 方程为

$$P'(t) = -\frac{G^2(t)}{D^2(t)}P^2(t).$$

其解为

$$P(t) = \left(\frac{1}{P(0)} + \int_0^t \left[\frac{G(s)}{D(s)}\right]^2 ds\right)^{-1}.$$

而 θ 的滤波过程（滤波估计）$\hat{\theta}_t$ 的 Kalman-Bucy 随机微分方程就是

$$\left(\frac{1}{P(0)} + \int_0^t \left[\frac{G(s)}{D(s)}\right]^2 ds\right)d\hat{\theta}_t = -\left[\frac{G(t)}{D(t)}\right]^2 \hat{\theta}_t dt + \frac{G(t)}{D(t)^2}d\eta_t,$$

它恰好就是

$$d\left(\left(\frac{1}{P(0)} + \int_0^t \left[\frac{G(s)}{D(s)}\right]^2 ds\right)\hat{\theta}_t\right) = \frac{G(t)}{D(t)^2}d\eta_t.$$

于是

$$\left(\frac{1}{P(0)} + \int_0^t \left[\frac{G(s)}{D(s)}\right]^2 ds\right)\hat{\theta}_t - \frac{1}{P(0)}\hat{\theta}_0 = \int_0^t \frac{G(s)}{D(s)^2}d\eta_s.$$

我们得到 θ 在时刻 t 的滤波估计 $\hat{\theta}_t$ 为

$$\hat{\theta}_t = \frac{\dfrac{\hat{\theta}_0}{P(0)} + \displaystyle\int_0^t \frac{G(s)}{D(s)^2}d\eta_s}{\dfrac{1}{P(0)} + \displaystyle\int_0^t \left[\frac{G(s)}{D(s)}\right]^2 ds}.$$

6.5.3　关于离散时间的 Kalman-Bucy 滤波的附注

（1）量测只依赖于同时刻的状态信息时的离散滤波模型

这是 Ito 过程形式的滤波模型的离散版本.

Kalman-Bucy 滤波模型是如下的一维的离散时间模型，即状态方程与量测方程分别为

$$\xi_{n+1} = F_n \xi_n + C_n u_n, \tag{6.40}$$
$$\eta_n = G_n \xi_n + D_n w_n, \tag{6.41}$$

其中$\{u_n\}$,$\{w_n\}$是相互独立的独立随机变量列,满足 $Eu_n = Ew_n = 0$,$\text{var}(u_n) = \text{var}(w_n) = 1$. 当 u_n, w_n 都是正态随机变量时,这相当于 Ito 过程形式的滤波模型的离散版本. 当 u_n, w_n 都未必是正态随机变量时,经典的考虑是线性滤波算法. 这时,相应于连续时间的 Kalman-Bucy 滤波随机微分方程的是如下的 Kalman-Bucy 迭代滤波公式

$$\hat{\xi}_{n+1} = F_n \hat{\xi}_n + K_n(\eta_n - G_n \hat{\xi}_n), \quad \hat{\xi}_0 = E\xi_0, \tag{6.42}$$

其中

$$K_n = \frac{G_n^2}{D_n^2} P_n \tag{6.43}$$

是增益系数,而

$$P_n = E(\xi_n - \hat{\xi}_n)^2$$

是滤波器在时刻 n 的均方误差,满足递推关系

$$P_n = \left[(F_{n-1}^2 P_{n-1} + C_{n-1}^2)^{-1} + \frac{G_{n-1}^2}{D_{n-1}^2}\right]^{-1}, \quad P(0) = \text{var}(\xi_0). \tag{6.44}$$

如果令

$$n_n \overset{\text{def}}{=\!=} \eta_n - G_n \hat{\xi}_n \quad (n_0 = \xi_0),$$

则 n_n 是**新息过程**,它满足

$$E(n_n n_m) = 0, \quad E(n_n \eta_m) = 0 \quad (m < n).$$

构造 Kalman-Bucy 滤波器的思想是,待定滤波随机序列$\{\hat{\xi}_n: n \geqslant 0\}$的随机差分方程(它是随机微分方程的离散版本)

$$\hat{\xi}_{n+1} = a_n \hat{\xi}_n + l_n n_n.$$

并归纳地确定待定的系数 l_n,使每一次更新的方差达到最小.

然而连续时间模型的更自然的离散版本,是如下的状态过程

$$\xi_n = \sum_{k=1}^{p} \alpha_k \xi_{n-k} + w_n, \tag{6.40$'$}$$

以及量测过程

$$\eta_t = \sum_{k=0}^{p-1} h_k \xi_{n-k} + v_n. \tag{6.41$'$}$$

在这里状态过程多个时刻的信息对量测有贡献,这样,用如下的多维离散滤波模型,可以更方便地发挥矩阵工具的长处.

(2) 量测依赖于多个时刻的状态信息时的常系数离散滤波模型

多个时刻的信息的滤波的数学处理,可以改记为如下的**多维 Kalman-Bucy 滤波模型**:

方程(6.40)′和方程(6.41)′可分别写成向量形式

$$\begin{pmatrix}\xi_n\\ \xi_{n-1}\\ \vdots\\ \xi_{n-p+1}\end{pmatrix}=\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0\end{pmatrix}\begin{pmatrix}\xi_{n-1}\\ \xi_{n-2}\\ \vdots\\ \xi_{n-p}\end{pmatrix}+\begin{pmatrix}w_n\\ 0\\ \vdots\\ 0\end{pmatrix},\tag{6.40$''$}$$

$$\eta_n=(h_0,h_1,\cdots,h_{p-1})\begin{pmatrix}\xi_n\\ \xi_{n-1}\\ \vdots\\ \xi_{n-p+1}\end{pmatrix}+v_n.\tag{6.41$''$}$$

我们将它们作一点推广,并加上干扰后,简记为

$$\boldsymbol{x}_n=\boldsymbol{F}\boldsymbol{x}_{n-1}+\boldsymbol{\Gamma}\boldsymbol{w}_n,\tag{6.40$'''$}$$

$$\boldsymbol{y}_n=\boldsymbol{H}\boldsymbol{x}_n+\boldsymbol{v}_n.\tag{6.41$'''$}$$

假定矩阵 $\boldsymbol{F}$ 在 $|z|\geqslant 1$ 中没有特征值. w_n,v_n,η_n 都是一维的随机变量,$\boldsymbol{\Gamma}$ 是 $p\times 1$ 矩阵,$\boldsymbol{H}$ 是 $1\times p$ 矩阵. 并将它们推广为:$\boldsymbol{x}_n$ 为 p 维向量值随机过程,$\boldsymbol{w}_n$ 为 r 维向量值随机过程,$\boldsymbol{y}_n,\boldsymbol{v}_n$ 为 m 维向量值随机过程,$\boldsymbol{F}$ 为 $p\times p$ 矩阵,$\boldsymbol{\Gamma}$ 为 $p\times r$ 矩阵,$\boldsymbol{H}$ 为 $m\times p$ 矩阵, 而 $\boldsymbol{w}_n,\boldsymbol{v}_n$ 都是均值为 0 的,相互独立的白噪声列(即独立同分布列),且相互独立,其方差矩阵分别记为

$$E(\boldsymbol{w}_n\boldsymbol{w}_n^{\mathrm{T}})=Q,\quad E(\boldsymbol{v}_n\boldsymbol{v}_n^{\mathrm{T}})=\boldsymbol{R}.\tag{6.45}$$

这里 $\{\boldsymbol{x}_n\}$ 是代表信号(随机的)状态列,而 $\{\boldsymbol{y}_n\}$ 是用来得到 $\{\boldsymbol{x}_n\}$ 的信息的(随机的)观测列,$\{\nu_n\}$ 是观测误差列. 这个模型称为一般的多维 Kalman-Bucy 滤波模型.

注　这样推广的向量模型的状态序列,包含了经典的 ARMA 时间序列,例如,ARMA(p,q)对应于

$$m=q+1,\quad \boldsymbol{w}_n=\begin{pmatrix}\varepsilon_n\\ \varepsilon_{n-1}\\ \vdots\\ \varepsilon_{n-q}\end{pmatrix},\quad \boldsymbol{\Gamma}=\begin{pmatrix}b_0 & b_1 & \cdots & b_q\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix},\quad \boldsymbol{Q}=\sigma^2\boldsymbol{I}_m,$$

其中 $\boldsymbol{I}_m$ 为表示 m 阶单位矩阵.

Kalman 和 Bucy 建立这个模型的目的,是用量测到的 $\boldsymbol{y}_n$ 来递推地估计测量不到的状态 $\boldsymbol{x}_n$. 我们将状态 $\boldsymbol{x}_n$ 的滤波估计记为 $\hat{\boldsymbol{x}}_n$. Kalman-Bucy 滤波的核心思想,就是要从观测列 $\{\boldsymbol{y}_n\}$ 出发,递推地对 $\{\boldsymbol{x}_n\}$ 作更新估计. 即利用对 $\boldsymbol{x}_{n-1}$ 的估计 $\hat{\boldsymbol{x}}_{n-1}$,及 $\boldsymbol{y}_n$ 得到 $\boldsymbol{x}_n$ 的估计 $\hat{\boldsymbol{x}}_n$. 其原则是:在已知 $\hat{\boldsymbol{x}}_{n-1}$ 时,选取矩阵 $\boldsymbol{A},\boldsymbol{B}$ 使方差矩阵

$$\boldsymbol{\Delta}_n\overset{\text{def}}{=\!=}E([\boldsymbol{x}_n-(\boldsymbol{A}\hat{\boldsymbol{x}}_{n-1}+\boldsymbol{B}\boldsymbol{y}_n)][\boldsymbol{x}_n-(\boldsymbol{A}\hat{\boldsymbol{x}}_{n-1}+\boldsymbol{B}\boldsymbol{y}_n)]^{\mathrm{T}})$$

最小,其中最小的含义应该理解成,在非负定对称矩阵的意义下的大小(即对非负对称矩阵 $\boldsymbol{C}_1,\boldsymbol{C}_2$,如果 $\boldsymbol{C}_1-\boldsymbol{C}_2$ 仍为非负定,则称为 $\boldsymbol{C}_1\geqslant\boldsymbol{C}_2$). 在确定了最佳的 $\boldsymbol{A},\boldsymbol{B}$ 以后,就取

$$\hat{\boldsymbol{x}}_n = \boldsymbol{A}\hat{\boldsymbol{x}}_{n-1} + \boldsymbol{B}\boldsymbol{y}_n.$$

注意到$\boldsymbol{\Delta}_n$ 是$\boldsymbol{A}$ 与$\boldsymbol{B}$ 的“二次式”，所以可以由“配平方”方法求得$\boldsymbol{\Delta}_n$ 的最小值. 这种做法的思想很简单(但是需要小心地注意不要把矩阵乘法的次序弄错)，写起来却有些繁琐，所以我们略去其推导而直接写出以下的结论.

定理 6.20　对于 Kalman-Bucy 滤波模型(6.40)‴，(6.41)‴，最佳线性滤波为

$$\hat{\boldsymbol{x}}_n = \boldsymbol{A}\hat{\boldsymbol{x}}_{n-1} + \boldsymbol{B}\boldsymbol{y}_n = \boldsymbol{F}\hat{\boldsymbol{x}}_{n-1} + \boldsymbol{K}_n(\boldsymbol{y}_n - \boldsymbol{H}\boldsymbol{F}\hat{\boldsymbol{x}}_{n-1}),$$

其中

$$\boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{B}\boldsymbol{H})\boldsymbol{F},\quad \boldsymbol{B} = \boldsymbol{\Pi}_n\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}_n\boldsymbol{\Pi}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R})^{-1} \stackrel{\text{def}}{=} \boldsymbol{K}_n \quad (\text{即增益系数}),$$

$$\boldsymbol{\Pi}_n = \boldsymbol{F}\boldsymbol{P}_{n-1}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}},\quad \boldsymbol{P}_n = E[(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n)(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n)^{\mathrm{T}}],$$

而$\boldsymbol{P}_n$ 是滤波的方差矩阵的估计.

时间序列的 ARMA 模型的预报问题，也可以作为 Kalman-Bucy 滤波的特殊情形，它对应与$m=1,H=1,v_n=0(R=0)$的情形.

然而，在实际操作时，需要知道模型参数$\boldsymbol{F},\boldsymbol{H},\boldsymbol{\Gamma},\boldsymbol{Q},\boldsymbol{R}$与初值$\boldsymbol{P}_0=\mathrm{var}(\boldsymbol{x}_0)$，$\hat{\boldsymbol{x}}_0=E\boldsymbol{x}_0$. $\boldsymbol{x}_0$ 可由我们对系统的认识来给定. 在 Kalman-Bucy 滤波的理论中有一个重要结论：在$\boldsymbol{F},\boldsymbol{H},\boldsymbol{\Gamma}$满足一个相当弱的条件(称为可观测条件)时，当$n\to+\infty$时，$\boldsymbol{P}_n$ 有一个与$\boldsymbol{P}_0$ 的取法无关的极限$\boldsymbol{P}$，$\boldsymbol{K}_n$ 也就应该趋向某个固定的$\boldsymbol{K}$. 也就是递推进行到适当步数后，$\boldsymbol{K}$ 就可以不必再变动了，这就会减少很多计算量.

注 1　(6.40)′式和(6.40)″式都可以对应地考虑系数显含时间t 的形式，也有相应的滤波均方误差矩阵序列的矩阵递推公式，以及滤波序列的递推公式.

注 2　连续时间的 Kalman-Bucy 滤波模型，也有其相应的多维 Ito 过程模型.

* 6.6　随机微分方程的弱解的概念

由推论 6.2 知道，在定理 6.1 的条件下，随机微分方程的解作为随机过程，可以抽象地用一个泛函关系

$$\xi_t = G(t,\xi_0,(B_s)_{0\leqslant s\leqslant t})$$

来表示. 在第 5 章中，我们已经知道，在随机微分方程中的 Brown 运动只是一个代表微观的随机涨落的驱动力，它并不在宏观的平均性质中出现. 所以说，它只是一个参变的 Brown 运动. 将随机微分方程中的 Brown 运动$\{B_t,t\geqslant 0\}$换成另一个 Brown 运动$\{\bar{B}_t,t\geqslant 0\}$时，其解就是$\bar{\xi}_t=G(t,\xi_0,(\bar{B}_s)_{0\leqslant s\leqslant t})$. 然而从分布(我们回忆随机过程的分布，粗略地，就是它的所有的有限维分布的全体)的角度看，随机过程$\{\xi_t,t\geqslant 0\}$的分布与随机过程$\{\bar{\xi}_t,t\geqslant 0\}$的分布是一样的. Stroock 与 Varadhan 在 20 世纪 70 年代初敏锐地注意到 Ito 公式存在一个用鞅叙述的方式(消去参变的 Brown 运动后)，它提供了随机微分方程(一个消去参变的

Brown 运动)用鞅叙述的方式,这就导出如下的定义.

定义 6.4 如果随机过程$\{\xi_t,t\geqslant 0\}$满足:对于一个对 x 二阶连续可微,且对 t 连续可微的任意函数 $f(t,x)$,如下定义的随机过程

$$M_t^{(f)} \overset{\text{def}}{=} f(t,\xi_t) - \int_0^t \left[f'_t(s,\xi_s) + f'_x(s,\xi_s)b(s,\xi_s) + \frac{1}{2}\sigma^2(s,\xi_s)f''_{xx}(s,\xi_s) \right] ds$$

都是鞅,那么,我们将随机过程$\{\xi_t,t\geqslant 0\}$称为**鞅问题的解**,或者称为随机微分方程(6.1)的**弱解**.

它之所以称为弱解,是因为此时在理论上可以证明,必定存在一个 Brown 运动$\{\bar{B}_t,t\geqslant 0\}$,使得随机过程$\{\xi_t,t\geqslant 0\}$是$(\bar{B}_t)$可知的,使得

$$\xi_t = \xi_0 + \int_0^t b(s,\xi_s)ds + \int_0^t \sigma(s,\xi_s)d\bar{B}_s.$$

因此,对于给定的 Brown 运动的随机微分方程的解,常常也称为**强解**. 强解与弱解的不同处仅仅在于,前者的 Brown 运动是给定的,而后者的 Brown 运动是找出来的. 由于物理系统所研究的是宏观的统计平均,所以,最重要的是确定这种系统(一般地是 Markov 系统)的转移密度,并不需要知道随机干扰是哪一个特定的 Brown 运动,而只需要有一个隐藏的如 Brown 运动那么快速变化的随机力干扰. 所以,在实用中常常只需要弱解的概念.

对于给定系数 $b(t,x)$,$\sigma(t,x)$的随机微分方程,为保证弱解的存在唯一性,对系数的要求可以大大地减弱. 例如,Stroock-Varadhan 定理指出,只要 $\sigma(t,x)$是有界连续函数,$b(t,x)$满足一致线性增长条件,即存在参数 C,使

$$|b(t,x)| \leqslant C(1+|x|),$$

那么弱解就存在唯一.

以数学角度看,弱解的存在唯一,对于所建立的模型的正确性,从数学理论的基础上给出了确保. 然而在实际的近似中,例如,用 **Monte Carlo 方法**(即**随机模拟方法**)近似计算,模拟系统的随机发展,或估计系统的未知参数时,仍旧需要考虑用强解的概念.

习题 6

1. 给出一般的形式随机微分方程

$$\frac{d^2\xi_t}{dt^2} = -\lambda^2\xi_t - b\frac{d\xi_t}{dt} + \sigma\dot{B}_t,$$

$$\xi_0, \quad \left.\frac{d\xi_t}{dt}\right|_{t=0} = v_0$$

的显式解.

2. 求 Langevin 方程的解的协方差函数 $\operatorname{cov}(\eta_t,\eta_s)$,并证明 $\lim\limits_{t-s\to 0}\operatorname{cov}(\eta_t,\eta_s)=0$.

3. 求证方程

$$d\xi_t = [c(t) + b(t)\xi_t]dt + \sum_{i=1}^{m}[\sigma_i(t) + a_i(t)\xi_t]dB_t^{(i)},$$
$$\xi_0 = \xi_0$$

的解是

$$\xi_t = F(t)\left[\xi_0 + \int_0^t F(s)^{-1}\left(c(s) - \sum_{i=1}^{m}\sigma_i(t)a_i(t)\right)ds + \int_0^t \sum_{i=1}^{m} F(s)^{-1}\sigma_i(s)dB_s^{(i)}\right],$$

其中

$$F(t) = e^{\int_0^t \left[b(s) - \frac{1}{2}\sum_{i=1}^{m} a_i^2(s)\right]ds + \int_0^t \sum_{i=1}^{m} a_i(s)dB_s^{(i)}}.$$

4. 求解 $d\xi_t = (a\xi_t + b\xi_t^\alpha)dt + c\xi_t dB_t$，其中 α 是一个实数.

5. 求解方程 $d\xi_t = -\frac{1}{2}e^{-2\xi_t}dt + e^{-\xi_t}dB_t$，并证明它在一个有限的随机时间爆炸.

6. 求解方程 $d\xi_t = -\xi_t dt + e^{-t}dB_t$.

7. 设 $\boldsymbol{B}_t = (B_t^{(1)}, B_t^{(2)}, \cdots, B_t^{(m)})$是 m 维 Brown 运动，求证 $R_t = |\boldsymbol{B}_t|$满足 Bessel 方程：

$$dR_t = \frac{m-1}{2R_t}dt + \sum_{i=1}^{m}\frac{B_t^{(i)}}{R_t}dB_t^{(i)}$$

8. 证明方程

$$d\xi_t = -\frac{1}{2}\xi_t dt - \eta_t dB_t,$$
$$d\eta_t = -\frac{1}{2}\eta_t dt + \xi_t dB_t$$

有解$(\xi_t, \eta_t) = (\cos B_t, \sin B_t)$，而且其任意一个解满足 $\sqrt{\xi_t^2 + \eta_t^2} = \sqrt{\xi_0^2 + \eta_0^2}$.

9. 求解方程

$$d\xi_t = dt + dB_t^{(1)},$$
$$d\eta_t = \xi_t dB_t^{(2)}.$$

10. 求解方程

$$d\xi_t = \eta_t dt + dB_t^{(1)},$$
$$d\eta_t = \xi_t dt + dB_t^{(2)}.$$

11. 对于光滑函数 $\sigma(x)$，证明 $g(B_t)$是方程

$$d\xi_t = \frac{1}{2}\sigma'(\xi_t)\sigma(\xi_t)dt + \sigma(\xi_t)dB_t,$$
$$\xi_0 = 0$$

的解，其中 $g(x)$ 是函数$\int_0^x \frac{du}{\sigma(u)}$ 的反函数.

12. 证明 $M_t = f(\boldsymbol{B}_t)e^{-\frac{1}{2}\int_0^t \Delta f(\boldsymbol{B}_s)ds}$ 是鞅，其中 $f(x)$ 是光滑函数，$\boldsymbol{B}_t$ 是m 维 Brown运动，其中 Δ 是 Laplace 算子.

第 7 章　扩散过程与其性质

7.1　随机微分方程解的 Markov 性质

作为随机微分方程的驱动的 Brown 运动，其轨道上在每一个时刻的变化是非常快的（即“速度”是无限的），在直观上可以想象，随机微分方程的解 ξ_t 的轨道也就变化得非常快，它类似于自然界的扩散现象. 所以导致以下的定义.

定义 7.1　随机微分方程

$$\mathrm{d}\xi_t = b(t,\xi_t)\mathrm{d}t + \sigma(t,\xi_t)\mathrm{d}B_t$$

的解作为随机过程，称为**扩散过程**，$b(t,x)$ 称为**漂移系数**，$a(t,x)\equiv\sigma^2(t,x)$ 称为**扩散系数**.

例 7.1（续例 6.11）　我们进一步分析例 6.11 中的随机微分方程的解

$$\xi_t = \mathrm{e}^{cB_t+\left(a-\frac{1}{2}c^2\right)t}\left[\sqrt{\xi_0}+\frac{b}{2}\int_0^t \mathrm{e}^{-\frac{1}{2}cB_s-\left(\frac{1}{4}c^2+\frac{a}{2}\right)s}\mathrm{d}s\right]^2.$$

这个解具有以下的典型性，正如在第 6 章中所示，随机过程 ξ_t 是 Brown 运动的泛函，而且是 (B_t) 可知的. 后者在直观上是显然的. 因为 Ito 方程的解代表了在 Brown 运动干扰下的物理系统的随机输出，它当然只依赖于干扰的历史情况，而决不会依赖于干扰的将来情况.

没有 Brown 运动干扰的 Ito 方程，就是普通的常微分方程，其解函数在给定时刻的值，只依赖于系统的初始值，而经 Brown 运动干扰下的 Ito 方程，其解作为随机过程在给定时刻就是一个随机变量，自然猜想在系统的初始值给定时，其分布应该是确定的. 让我们通过例 6.11 验证这个想法. 对此，由

$$\xi_t = \mathrm{e}^{cB_t+\left(a-\frac{1}{2}c^2\right)t}\left[\sqrt{\xi_0}+\frac{b}{2}\int_0^t \mathrm{e}^{-\frac{1}{2}cB_s-\left(\frac{1}{4}c^2+\frac{a}{2}\right)s}\mathrm{d}s\right]^2,$$

得到

$$\xi_{t+s} = \mathrm{e}^{cB_t+\left(a-\frac{1}{2}c^2\right)t}\left[\sqrt{\xi_s}+\frac{b}{2}\int_0^t \mathrm{e}^{-\frac{1}{2}c(B_{s+u}-B_s)-\left(\frac{1}{4}c^2+\frac{a}{2}\right)u}\mathrm{d}u\right]^2.$$

注意到当 $u\leqslant s$ 时，ξ_u 仅依赖于 $\{B_v, v\leqslant u\}$，而且 ξ_{t+s} 的表达式中所含的随机部分是由两部分组成：第一部分是 ξ_s，第二部分是 Brown 运动在时刻 s 后的增量，所以 ξ_{t+s} 在条件 $\xi_s=x$ 下的条件分布是与 $\{B_v, v\leqslant u\}$（其中 $u\leqslant s$）独立的. 因此其条件分布函数

$$
\begin{aligned}
&P(\xi_{t+s}\leqslant y\mid \xi_s=x,\xi_u=x_u(0\leqslant u<s))\\
&=P\left(\left[\mathrm{e}^{\frac{c}{2}(B_{t+s}-B_s)+\left(\frac{1}{4}c^2-\frac{a}{2}\right)t}x+\frac{b}{2}\int_s^{t+s}\mathrm{e}^{\frac{c}{2}(B_{t+s}-B_v)-\left(\frac{1}{4}c^2-\frac{a}{2}\right)(t+s-v)}\mathrm{d}v\right]^2\leqslant y\right)\\
&=P(\xi_{t+s}\leqslant y\mid \xi_s=x).
\end{aligned}
$$

另一方面，Brown 运动有平移的统计不变性质，随机过程 $\{B_{t+s}-B_v,s\leqslant v\leqslant s+t\}$ 与随机过程 $\{B_t-B_{v-s},0\leqslant v-s\leqslant t\}$ 有相同的分布（即所有的有限维分布都相同，而且经相同的运算，或相同的积分以后的分布也相同），也就是

$$
\mathrm{e}^{\frac{c}{2}(B_{t+s}-B_s)+\left(\frac{1}{4}c^2-\frac{a}{2}\right)t}\quad 与\quad \mathrm{e}^{\frac{c}{2}B_t+\left(\frac{1}{4}c^2-\frac{a}{2}\right)t}\ 同分布
$$

$$
\int_s^{t+s}\mathrm{e}^{\frac{c}{2}(B_{t+s}-B_v)-\left(\frac{1}{4}c^2-\frac{a}{2}\right)(t+s-v)}\mathrm{d}v\quad 与\quad \int_0^t\mathrm{e}^{\frac{c}{2}(B_s-B_r)-\left(\frac{1}{4}c^2-\frac{a}{2}\right)(t-r)}\mathrm{d}r\ 同分布，
$$

故而

$$
\begin{aligned}
&P(\xi_{t+s}\leqslant y\mid \xi_s=x)\\
&=P\left([\mathrm{e}^{\frac{c}{2}B_t+\left(\frac{1}{4}c^2-\frac{a}{2}\right)t}x+\frac{b}{2}\int_0^t\mathrm{e}^{\frac{c}{2}(B_t-B_r)-\left(\frac{1}{4}c^2-\frac{a}{2}\right)(t-r)}\mathrm{d}v]^2\leqslant y\right)=P(\xi_t\leqslant y\mid \xi_0=x).
\end{aligned}
$$

这样我们就得到了如下的结论：这个随机微分方程的解是时齐的 Markov 过程.

这个结论可以一般化，即由定义 7.1 定义的任意随机微分方程的解，都是 Markov 过程. 事实上，上面的证明已经概括了全貌与实质. 而在一般情形下的想法与推导，反而因为摆脱了具体例子中出现的复杂演算，而更突出其本质，这就是数学抽象的简捷性与长处.

定理 7.1　随机微分方程的解是 Markov 过程. 而当系数不含时间 t 时，即 $\sigma(t,x)=\sigma(x)$，$b(t,x)=b(x)$时，此 Markov 过程是时齐的.

证明　对于任意 n 及任意 $t>s>s_n>\cdots>s_1$，在 $t\in[s,+\infty)$，满足条件 $\xi_s=x,\xi_{s_n}=x_n,\cdots,\xi_{s_1}=x_1$ 的解 ξ_t 的迭代构造也应为

$$
\xi_t^{(0)}=x,
$$

$$
\xi_t^{(n+1)}=x+\int_s^t b(u,\xi_u^{(n)})\mathrm{d}u+\int_s^t\sigma(u,\xi_u^{(n)})\mathrm{d}B_u,
$$

其中的一切 $\xi_t^{(n)}$ 都只依赖于 $u,x,\sigma(u,x),b(u,x)$，而不依赖于 $x_n,\cdots,x_1$. 从而条件概率 $P(\xi_t\in\Lambda\mid\xi_s=x,\xi_{s_n}=x_n,\cdots,\xi_{s_1}=x_1)$也不依赖于 $x_n,\cdots,x_1$，即应该有

$$
P(\xi_t\in\Lambda\mid\xi_s=x,\xi_{s_n}=x_n,\cdots,\xi_{s_1}=x_1)=P(\xi_t\in\Lambda\mid\xi_s=x),
$$

当系数不显含 t 时，即 $\sigma(t,x)=\sigma(x)$，$b(t,x)=b(x)$时，对于方程的解有

$$
\xi_{t+s}=\xi_s+\int_s^{t+s}b(\xi_u)\mathrm{d}u+\int_s^{t+s}\sigma(\xi_u)\mathrm{d}B_u.
$$

令

$$
\bar{\xi}_t\overset{\mathrm{def}}{=\!=}\xi_{t+s},\quad \bar{B}_t\overset{\mathrm{def}}{=\!=}B_{t+s}-B_s,
$$

则 $\bar{B}_t$ 也是一个 Brown 运动. 而且直接由 Ito 积分的定义,取极限容易得到

$$\int_s^{t+s}\sigma(\xi_u)\mathrm{d}B_u = \int_0^t \sigma(\bar{\xi}_u)\mathrm{d}\bar{B}_u.$$

于是随机微分方程

$$\bar{\xi}_t = \bar{\xi}_0 + \int_0^t b(\bar{\xi}_u)\mathrm{d}u + \int_0^t \sigma(\bar{\xi}_u)\mathrm{d}\bar{B}_u$$

与随机微分方程

$$\xi_t = \xi_0 + \int_0^t b(\xi_u)\mathrm{d}u + \int_0^t \sigma(\xi_u)\mathrm{d}B_u$$

是具有相同系数,但对应于不同 Brown 运动的 Ito 方程的解. 由于对于不同的 Brown 运动对应的迭代过程的每一步都具有相同的分布,因此,两个随机微分方程的解作为迭代过程的极限,也就有相同的转移函数. 这就证明了这个 Markov 过程是时齐的.

注　系数只满足局部一致 Lipshitz 条件的一般的随机微分方程的解,也是 Markov 过程,但是,这时对随机微分方程的解的定义,需要稍作一些修正,因为这时的解有可能发生"爆炸",即解在一个有限的时刻可能取 $-\infty$ 或 $+\infty$ $\Big($这个时刻还依赖于轨道,即是随机的时刻. 类似的情形也发生在常微分方程,例如,$\mathrm{d}x_t = x_t^2\mathrm{d}t$,即 $x' = x^2$,这时的解是 $x(t) = -\dfrac{1}{t-c}$,有 $\lim\limits_{t\uparrow c} x(t) = +\infty$. 即 $x(t)$ 在时刻 c"爆炸". 不同的是,这里没有作为随机驱动力的 Brown 运动,故而只有一条确定的轨道 $x(t) = -\dfrac{1}{t-c}$,因此爆炸时刻不是随机的$\Big)$. 这时如果在 $\xi_s = x$ 的条件下,对于 $t > s$,ξ_t 有条件密度,记为 $p(s,x,t,y)$(称为转移密度),将是"亏缺的"密度,即只能保证 $\int p(s,x,t,y)\mathrm{d}y \leqslant 1$,而不一定有 $\int p(s,x,t,y)\mathrm{d}y = 1$. 直观地可以理解为:有一部分概率消失了.

例 7.2(OU 过程的转移密度)　例 6.1 中的 Langevin 方程

$$\mathrm{d}\eta_t = -b\eta_t\mathrm{d}t + \sigma\mathrm{d}B_t$$

的解为

$$\eta_t = \mathrm{e}^{-bt}\left[\eta_0 + \sigma\int_0^t \mathrm{e}^{bu}\mathrm{d}B_u\right],$$

称为 **OU 过程**(**Ornstein-Ulenbeck 过程**). 对于 $t > s$,将 η_{s+t} 与

$$\eta_s = \mathrm{e}^{-bs}\left[\eta_0 + \sigma\int_0^s \mathrm{e}^{bu}\mathrm{d}B_u\right]$$

比较后,得到

$$\eta_{s+t} = \mathrm{e}^{-bt}\left[\eta_s + \sigma\int_s^{s+t}\mathrm{e}^{b(u-s)}\mathrm{d}B_u\right].$$

由第 5 章中,我们证明了

$$\int_s^{s+t} e^{bu} dB_u \sim N\left(0, \int_s^{s+t} e^{2bu} du\right) = N\left(0, \frac{e^{2b(s+t)} - e^{2bs}}{2b}\right).$$

故而,在 $\eta_s = x$ 的条件下,η_{s+t}的条件分布是正态分布 $N\left(xe^{-bt}, \frac{\sigma^2}{2b}(1-e^{-2bt})\right)$.

于是,Langevin 方程的转移密度为

$$p(t,x,y) = \sqrt{\frac{b}{\pi\sigma^2(1-e^{-2bt})}} \exp\left(-\frac{(y - xe^{-bt})^2}{\frac{\sigma^2(1-e^{-2bt})}{b}}\right).$$

例 7.3(Black-Scholes 随机微分方程的转移密度) 对于 Black-Scholes 随机微分方程

$$d\xi_t = \xi_t(bdt + \sigma dB_t)$$

的解

$$\xi_t \overset{\text{def}}{=\!=} \xi_0 e^{\left(b-\frac{\sigma^2}{2}\right)t+\sigma B_t}$$

有

$$\begin{aligned} F(t,x,y) &= P(\xi_{t+s} \leqslant y \mid \xi_s = x) \\ &= P(\xi_s e^{\left(b-\frac{\sigma^2}{2}\right)t+\sigma(B_{t+s}-B_s)} \leqslant y \mid \xi_s = x) \\ &= P(xe^{\left(b-\frac{\sigma^2}{2}\right)t+\sigma(B_{t+s}-B_s)} \leqslant y) \\ &= P\left(\frac{B_{t+s}-B_s}{\sqrt{t}} \leqslant \frac{1}{\sigma\sqrt{t}}\left[\ln\frac{y}{x} - \left(b-\frac{\sigma^2}{2}\right)t\right]\right) \\ &= \Phi\left(\frac{1}{\sigma\sqrt{t}}\left[\ln\frac{y}{x} - \left(b-\frac{\sigma^2}{2}\right)t\right]\right), \end{aligned}$$

其中 Φ 是标准正态分布函数.

命题 7.2 OU 过程是具有转移密度的 Markov 过程,当 $t \to +\infty$时,其转移密度有极限

$$\lim_{t\to+\infty} p(t,x,y) = \frac{1}{\sigma\sqrt{\frac{\pi}{b}}} e^{-\frac{y^2}{\frac{\sigma^2}{b}}} \overset{\text{def}}{=\!=} \pi(y).$$

即其极限分布为 $N\left(0, \frac{\sigma^2}{2b}\right)$,它不依赖于初始状态 x.

例 7.4(OU 过程的 Gauss 性) 在初值 $\eta_0 = x$ 条件下的 OU 过程,在 $t>0$ 时,是 Gauss 过程,且满足

$$E\eta_t = xe^{-bt}.$$

对于 $0<s<t$ 有

$$\mathrm{cov}(\eta_t,\eta_s)=\frac{\sigma^2}{2b}(\mathrm{e}^{-b(t-s)}-\mathrm{e}^{-b(t+s)}),$$

$$\mathrm{var}(\eta_t)=\frac{\sigma^2}{2b}(1-\mathrm{e}^{-2bt}).$$

当初值 $\eta_0\sim N(\mu_0,\sigma_0^2)$，且与整个 Brown 运动独立时，OU 过程在 $t\geqslant 0$ 时也是 Gauss 过程，这时

$$E\eta_t=(E\eta_0)\mathrm{e}^{-bt},$$

对于 $0<s<t$ 有

$$\mathrm{cov}(\eta_t,\eta_s)=\sigma_0^2\mathrm{e}^{-b(t+s)}+\frac{\sigma^2}{2b}(\mathrm{e}^{-b(t-s)}-\mathrm{e}^{-b(t+s)}),$$

$$\mathrm{var}(\eta_t)=\sigma_0^2\mathrm{e}^{-2bt}+\frac{\sigma^2}{2b}(1-\mathrm{e}^{-2bt}).$$

特别地，如果取初值为极限分布，即 $\eta_0\sim N\left(0,\frac{\sigma^2}{2b}\right)$，那么 OU 过程就成为**平稳过程**，即对于任意 n 及任意 $0\leqslant t_1<\cdots<t_n,s>0$，$(\xi_{t_1+s},\xi_{t_2+s},\cdots,\xi_{t_n+s})$与$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$有相同的联合密度，而且

$$E\eta_t\equiv 0,\quad \mathrm{cov}(\eta_t,\eta_s)=\frac{\sigma^2}{2b}\mathrm{e}^{-b(t-s)}.$$

因此

$$\mathrm{var}(\eta_t)=\frac{\sigma^2}{2b}.$$

其验证如下：

首先它满足 Langevin 方程，由此直接推出它是 Gauss 过程. 其次，我们计算它的协方差函数. 当 $\eta_0=x$ 时，对于 $0<s<t$ 有

$$\begin{aligned}\mathrm{cov}(\eta_t,\eta_s)&=E\left[\mathrm{e}^{-bt}\sigma\int_0^t\mathrm{e}^{bu}\mathrm{d}B_u\mathrm{e}^{-bs}\sigma\int_0^s\mathrm{e}^{bu}\mathrm{d}B_u\right]\\&=\sigma^2\mathrm{e}^{-b(t+s)}\int_0^s\mathrm{e}^{2bu}\mathrm{d}u=\frac{\sigma^2}{2b}(\mathrm{e}^{-b(t-s)}-\mathrm{e}^{-b(t+s)}).\end{aligned}$$

而当 $\eta_0\sim N(\mu_0,\sigma_0^2)$时，利用协方差运算的双线性性质，得到

$$\begin{aligned}\mathrm{cov}(\eta_t,\eta_s)&=\mathrm{cov}\left(\mathrm{e}^{-bt}\eta_0+\mathrm{e}^{-bt}\sigma\int_0^t\mathrm{e}^{bu}\mathrm{d}B_u,\mathrm{e}^{-bs}\eta_0+\mathrm{e}^{-bs}\sigma\int_0^s\mathrm{e}^{bu}\mathrm{d}B_u\right)\\&=\sigma_0^2\mathrm{e}^{-b(t+s)}+\frac{\sigma^2}{2b}(\mathrm{e}^{-b(t-s)}-\mathrm{e}^{-b(t+s)}).\end{aligned}$$

最后，在 $\eta_0\sim N\left(0,\frac{\sigma^2}{2b}\right)$时，OU 过程是平稳过程的证明，这是因为

$$\mathrm{cov}(\eta_t,\eta_s)=\frac{\sigma^2}{2b}\mathrm{e}^{-b(t-s)},$$

$$\mathrm{var}(\eta_t - \eta_s) = \frac{\sigma^2}{2b}\mathrm{e}^{-b(t-s)}.$$

7.2 扩散方程与 Fokker-Planck 方程

7.2.1 转移密度的 Kolmogorov 向前方程与向后方程

与 Langevin 方程有转移密度类似，对于相当一般的随机微分方程的解作为随机过程都存在转移密度. 对此我们不加证明地引用如下结论.

定理 7.3 相当一般(例如，处理定理 6.1 的条件外，还满足：系数光滑，扩散系数有一个正的下界的情形)的随机微分方程的解，是具有转移密度 $p(s,x,t,y)$的 Markov 过程.

在系数不含 t 时，如果 $\sigma(x)$，$b(x)$有相当次数的连续导数，可以证明线性偏微分方程

$$\frac{\partial u(t,x)}{\partial t} = \frac{1}{2}\sigma^2(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x}$$

的基本解就是转移密度 $p(t,x,y) \equiv p(0,x,t,y)$. 即对于任意固定的 x，$p(t,x,y)$满足

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2(x)\frac{\partial^2 p}{\partial x^2} + b(x)\frac{\partial p}{\partial x} \quad \left(\text{简记为}\frac{\partial p}{\partial t} = Lp\right),$$

$$\int p(0,x,y)f(y)\mathrm{d}y = f(x) \quad (\text{初值}).$$

这个转移密度 $p(t,x,y)$满足的线性偏微分方程，称为**扩散方程**，也称 **Kolmogorov 向后方程**，对应于扩散系数 $\sigma^2(x)$和漂移系数 $b(x)$. 由方程

$$\xi_t - \xi_s = \int_s^t b(\xi_u)\mathrm{d}u + \int_s^t \sigma(\xi_u)\mathrm{d}B_u,$$

可以直接看出，漂移系数与扩散系数的概率含义为

$$b(x) = \lim_{h\to 0}\frac{E[(\xi_{t+h} - \xi_t) \mid \xi_t = x]}{h} \quad (\textbf{即条件平均速率}),$$

$$\sigma^2(x) = \lim_{h\to 0}\frac{E[(\xi_{t+h} - \xi_t)^2 \mid \xi_t = x]}{h} \quad (\textbf{即条件平均二阶矩增长率}).$$

再则

$$E[f(\xi_t) \mid \xi_0 = x] = \int p(t,x,y)f(y)\mathrm{d}y.$$

注 转移密度的存在性的证明较为复杂，需要借助于较为深入的数学工具，如线性偏微分方程的基本解方法，或借助于泛函分析中广义解的 Weyl 引理，或借助于随机变分学中的 Malliavin 导数. 本书不想涉及这些材料.

在转移密度存在的前提下，我们可以得到 Kolmogorov 向后方程的直观推导，这将在下面给出.

7.2.2 转移密度的 Kolmogorov 向前方程与向后方程的直观推导

定理 7.4(Kolmogorov 向前方程,Fokker-Planck 方程) 在定理 7.3 同样的假定下且系数不含 t 的时齐情形,扩散过程的转移密度 $p(t,x,y)$是方程

$$\frac{\partial p}{\partial t}=\frac{\partial^2}{\partial y^2}\left(\frac{\sigma(y)^2}{2}p\right)-\frac{\partial}{\partial y}(b(y)p) \quad \left(\text{简记为}\frac{\partial p}{\partial t}=L^* p\right)$$

的基本解,即它还满足

$$\int p(0,x,y)f(y)\mathrm{d}y=f(x) \quad (\text{初值}).$$

这个方程称为 **Kolmogorov 向前方程**,或 **Fokker-Planck 方程**. 线性微分运算 L^* 称为线性微分运算 L 的**形式共轭运算**,其含义为对于任意 C_0^2(紧支集的(即只在有限区间上可能取非零值的)二阶连续可微函数类)函数 f 与 g,恒有

$$\int(Lf)(x)g(x)\mathrm{d}x=\int f(x)(L^* g)(x)\mathrm{d}x.$$

推论 7.5 对于有界连续函数 $f(x)$,函数

$$u(t,x)=E_x f(\xi_t)\overset{\text{def}}{=\!=}E[f(\xi_t)\mid\xi_0=x]$$

是方程

$$\begin{cases}\dfrac{\partial u}{\partial t}=\dfrac{1}{2}\sigma^2(x)\dfrac{\partial^2 u}{\partial x^2}+b(x)\dfrac{\partial u}{\partial x},\\ u(0,x)=f(x)\end{cases}$$

的解.

形式证明 由于 $u(x)=E[f(\xi_t)|\xi_s=x]=\int p(t,x,y)f(y)\mathrm{d}y$,得到

$$\begin{aligned}\frac{\partial u(t,x)}{\partial t}&=\int\frac{\partial p(t,x,y)}{\partial t}f(y)\mathrm{d}y=\int(L^* p)(t,x,y)f(y)\mathrm{d}y\\&=\int\left[\frac{1}{2}\sigma^2(x)f''(y)+b(x)f'(y)\right]p(t,x,y)\mathrm{d}y\\&=\frac{1}{2}\sigma^2(x)\frac{\partial^2 u}{\partial x^2}+b(x)\frac{\partial u}{\partial x}\qquad(\text{积分与微商交换次序}),\end{aligned}$$

其中 L^* 表示微分运算 L 的形式共轭运算.

这个线性微分方程的解的这个概率表示的公式,称为 **Kolmogorov 公式**,用它可以对于解 $u(t,x)$的值作随机模拟估计.

Fokker-Planck 方程的另一种更为常见的形式,是**散度型**的方程,就是

$$\frac{\partial p}{\partial t}=\frac{\partial}{\partial y}\left(\frac{\sigma(y)^2}{2}\frac{\partial p}{\partial y}\right)-\frac{\partial}{\partial y}([b(y)+\sigma(y)\sigma'(y)]p).$$

Kolmogorov 向前方程与向后方程的直观推导如下:

(1) Kolmogorov 向前方程的直观推导. 由 Ito 公式有

$$\mathrm{d}f(\xi_t) = f'(\xi_t)\mathrm{d}\xi_t + \frac{1}{2}f''(\xi_t)(\mathrm{d}\xi_t)^2,$$

$$f(\xi_{t+s}) = f(\xi_s) + \int_s^{s+t}\left[\frac{1}{2}\sigma^2(\xi_u)f''(\xi_u) + b(\xi_u)f'(\xi_u)\right]\mathrm{d}t + \int_s^{s+t}\sigma(\xi_u)f'(\xi_u)\mathrm{d}B_t.$$

对于有连续二阶导数且在某个有限区间外取 0 值的函数(简称为 C_0^2 函数)$f(x)$而言, $\int_0^t b(\xi_u)f(\xi_u)\mathrm{d}B_u$ 是(B_t)鞅. 记

$$E_x\eta = E(\eta \mid \xi_0 = x),\quad \mathcal{F}_t = \{\xi_u: u \leqslant t\},\quad \boldsymbol{B}_t = \{B_u: u \leqslant t\}.$$

由条件期望的性质及 Markov 性, 得到

$$E_x[f(\xi_{t+s}) - E(f(\xi_t))] = E_xE_x(E_x[(f(\xi_{s+t}) - f(\xi_t)) \mid \boldsymbol{B}_t] \mid \mathcal{F}_t)$$

$$= E_xE_x\left(E_x\left(\int_t^{s+t}\left[\frac{1}{2}\sigma^2(\xi_u)f''(\xi_u) + b(\xi_u)f'(\xi_u)\right]\mathrm{d}u\right)\middle|\boldsymbol{B}_t\right]\middle|\mathcal{F}_t\right)$$

$$= E_xE_x\left(\int_t^{s+t}\left[\frac{1}{2}\sigma^2(\xi_u)f''(\xi_u) + b(\xi_u)f'(\xi_u)\right]\mathrm{d}u\middle|\mathcal{F}_t\right)$$

$$= E_xE_x\left(\int_t^{s+t}\left[\frac{1}{2}\sigma^2(\xi_u)f''(\xi_u) + b(\xi_u)f'(\xi_u)\right]\mathrm{d}u\middle|\xi_t\right).$$

用微积分中的定理导致

$$\frac{1}{s}E_x[f(\xi_{t+s}) - f(\xi_t)] = E_x\left(\frac{1}{s}\int_t^{s+t}E_x\left[\left(\frac{1}{2}\sigma^2(\xi_u)f''(\xi_u) + b(\xi_u)f'(\xi_u)\right)\middle|\xi_t\right]\mathrm{d}u\right)$$

$$\xrightarrow{s\to 0} E\left[\left(\frac{1}{2}\sigma^2(\xi_t)f''(\xi_t) + b(\xi_t)f'(\xi_t)\right) \mid \xi_0 = x\right].$$

另一方面, 如果假定时齐的 Markov 过程 ξ_t 有转移密度 $p(t,x,y)$, 那么由

$$E(f(\xi_t) \mid \xi_0 = x) = \int p(t,x,y)f(y)\mathrm{d}y,$$

上面的极限式就变成

$$\frac{1}{s}\int(p(t+s,x,y) - p(t,x,y))f(y)\mathrm{d}y$$

$$\xrightarrow{s\to 0}\int p(t,x,y)\left[\frac{1}{2}\sigma^2(y)f''(y) + b(y)f'(y)\right]\mathrm{d}y.$$

也就是

$$\int\frac{\partial p(t,x,y)}{\partial t}f(y)\mathrm{d}y = \int p(t,x,y)\left[\frac{1}{2}\sigma^2(y)f''(y) + b(y)f'(y)\right]\mathrm{d}y.$$

再用分部积分推出(利用 f 在$\pm\infty$处为 0)

$$\int\frac{\partial p(t,x,y)}{\partial t}f(y)\mathrm{d}y = \int\left[\frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y)p(t,x,y)) - \frac{\partial}{\partial y}(b(y)p(t,x,y))\right]f(y)\mathrm{d}y.$$

由于 f 的任意性, 我们便得到

$$\frac{\partial p(t,x,y)}{\partial t}=\frac{1}{2}\frac{\partial^2}{\partial y^2}\Big(\sigma^2(y)p(t,x,y)\Big)-\frac{\partial}{\partial y}\Big(b(y)p(t,x,y)\Big).$$

这就得到 Kolmogorov 向前方程.

(2) Kolmogorov 向后方程的直观推导. 对于任意 C_0^2 函数 $g(x)$,令

$$v(t,y)=\int p(t,x,y)g(x)\mathrm{d}x.$$

那么,由 Kolmogorov 向前方程有

$$\frac{\partial v(t,y)}{\partial t}=\frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y)v(t,y))-\frac{\partial}{\partial y}(b(y)v(t,y)).$$

另一方面,利用 Chapman-Kolmogorov 方程,还有

$$\begin{aligned}v(t+s,y)&=\int p(t+s,x,y)g(x)\mathrm{d}x\\&=\iint p(t,x,z)p(s,z,y)\mathrm{d}zg(x)\mathrm{d}x\\&=\int p(s,z,y)v(t,z)\mathrm{d}z=\int p(s,x,y)v(t,x)\mathrm{d}x\\&\quad\left(\text{对称地得}\int p(t,x,y)v(s,x)\mathrm{d}x\right)\end{aligned}$$

于是,从此式及对 Kolmogorov 向前方程作"形式的"分部积分(终端 $\pm\infty$ 处都"认为"是 0),推出

$$\begin{aligned}\int\frac{\partial p(t+s,x,y)}{\partial t}g(x)\mathrm{d}x&=\frac{\partial v(t+s,y)}{\partial t}=\frac{\partial v(t+s,y)}{\partial s}\\&=\int p(t,x,y)\frac{\partial v(s,x)}{\partial s}\mathrm{d}x\\&=\int p(t,x,y)\left(\frac{1}{2}\frac{\partial^2[\sigma^2(x)v(s,x)]}{\partial x^2}-\frac{\partial[b(x)v(s,x)]}{\partial x}\right)\mathrm{d}x\\&=\int v(s,x)\left[\frac{1}{2}\sigma^2(x)\frac{\partial^2 p(t,x,y)}{\partial x^2}+b(x)\frac{\partial p(t,x,y)}{\partial x}\right]\mathrm{d}x\\&=\iint p(s,y,x)g(y)\mathrm{d}y\left[\frac{1}{2}\sigma^2(x)\frac{\partial^2 p(t,x,y)}{\partial x^2}+b(x)\frac{\partial p(t,x,y)}{\partial x}\right]\mathrm{d}x.\end{aligned}$$

令 $s\to 0$ 就得到

$$\int\frac{\partial p(t,x,y)}{\partial t}g(x)\mathrm{d}x=\int g(x)\left[\frac{1}{2}\sigma^2(x)\frac{\partial^2 p(t,x,y)}{\partial x^2}+b(x)\frac{\partial p(t,x,y)}{\partial x}\right]\mathrm{d}x.$$

再由 $g(x)$ 的任意性就得到

$$\frac{\partial p(t,x,y)}{\partial t}=\frac{1}{2}\sigma^2(x)\frac{\partial^2 p(t,x,y)}{\partial x^2}+b(x)\frac{\partial p(t,x,y)}{\partial x}.$$

这就形式地导出了 Kolmogorov 向后方程.

需要指出的是，以上均为形式运算，对于函数的定义域，对于积分、微商、极限等诸多运算之间交换的合理性，都未于追究．事实上还需要作一些假定（而且以上的直观证明是完全基于存在转移密度的假定下的）．

注　转移密度的统计估计，可以通过模拟随机过程的多个轨道，用统计中的直方图，或较为精确一些的核估计方法粗略地近似．

7.2.3　系数含时间 t 时的 Kolmogorov 向前方程与向后方程（非时齐形式）

定理 7.4′　在随机微分方程的系数含 t 时，其解 $\{\xi_t, t\geqslant 0\}$ 是具有转移密度 $p(s,x,t,y)$（ξ_t 在条件 $\xi_s=x$ 下的条件概率分布的密度函数在 y 处的值，$s<t$）的非时齐 Markov 过程．当 $\sigma(t,x)$，$b(t,x)$ 有相当次数的连续导数时，可以证明转移密度 $p(s,x,t,y)$ 满足

$$-\frac{\partial p}{\partial s}=\frac{1}{2}\sigma^2(s,x)\frac{\partial^2 p}{\partial x^2}+b(s,x)\frac{\partial p}{\partial x},$$

$$\lim_{s\uparrow t}\int p(s,x,t,y)f(y)\mathrm{d}y=f(x).$$

称为 **Kolmogorov 向后方程**（对于比时间 t 小的，落后的时间 s，及所在的状态 x），其中 $\sigma^2(t,x)$ 与 $b(t,x)$ 分别地是扩散系数与漂移系数，而其概率含义仍然是（不同处仅在于，这里的系数显含时间 t）

$$b(t,x)=\lim_{h\to 0}\frac{E[(\xi_{t+h}-\xi_s)\mid \xi_t=x]}{h}\quad（条件平均速率），$$

$$\sigma^2(t,x)=\lim_{h\to 0}\frac{E[(\xi_{t+h}-\xi_i)^2\mid \xi_t=x]}{h}\quad（条件平均二阶矩增长率）．$$

转移密度 $p(s,x,t,y)$ 还满足如下的方程

$$\frac{\partial p}{\partial t}=\frac{\partial^2}{\partial y^2}\left(\frac{\sigma(t,y)^2}{2}p\right)-\frac{\partial}{\partial y}\left(b(t,y)p\right),$$

$$\lim_{t\downarrow s}\int p(s,x,t,y)f(y)\mathrm{d}y=f(x).$$

它称为 **Kolmogorov 向前方程**（对于比时间 s 大的，超前的时间 t，及所在的状态 y），或 **Fokker-Planck 方程**．

相应于推论 7.5 的非时齐情形是下面的结论．

推论 7.5′　对于有界连续函数 $f(x)$，函数

$$u(t,x)\overset{\text{def}}{=\!=}E[f(\xi_T)\mid \xi_t=x]\quad(t\leqslant T)$$

是方程

$$\begin{cases}\dfrac{\partial u}{\partial t}+\dfrac{1}{2}\sigma^2(t,x)\dfrac{\partial^2 u}{\partial x^2}+b(t,x)\dfrac{\partial u}{\partial x}=0,\\ \lim\limits_{t\uparrow T}u(t,x)=f(x)\end{cases}$$

的解.

其形式证明与推论 7.5 的形式证明十分类似. 而推论 7.5 是推论 7.5′的特殊情形，在那里的转移密度 $p(s,x,t,y)$是时齐的，即 $p(s,x,t,y)\equiv p(0,x,t-s,y)\overset{\text{def}}{=\!=}p(t-s,x,y)$. 只要在推论 7.5′中令 $t=T-s$，并记

$$v(t,x)\equiv u(T-t,x)=E[f(\xi_T)\mid \xi_{T-t}=x]=E[f(\xi_t)\mid \xi_0=x]=E_x f(\xi_t),$$

就得到对于 $t\leqslant T$ 有

$$\begin{cases}\dfrac{\partial v}{\partial t}=\dfrac{1}{2}\sigma^2(x)\dfrac{\partial^2 v}{\partial x^2}+b(x)\dfrac{\partial v}{\partial x},\\ v(0,x)=f(x)\end{cases}$$

但是 T 可以任意取，这就得到推论 7.5 中的 Kolmogorov 概率表示公式.

定理 7.6(Master 方程)　假设初值 ξ_0 有分布密度(此条件并不必要，是为了数学处理更简单而作的冗余假定). 在定理 7.3 的假定下，将随机微分方程的解$\{\xi_t,t\geqslant 0\}$在时刻 t 的分布函数 $P(\xi_t\leqslant y)$的密度函数(正如前面指出的，它必定存在，但是却不容易证明)记为 $p(t,y)$，那么，它满足以下的 **Master 方程**：

$$\frac{\partial p}{\partial t}=\frac{\partial^2}{\partial y^2}\left(\frac{\sigma(y)^2}{2}p\right)-\frac{\partial}{\partial y}(b(y)p).$$

证明　记 ξ_0 的密度为 $p(0,y)$. 由于 $p(t,y)=\int p(t,x,y)p(0,x)\mathrm{d}x$，将转移密度满足的 Kolmogorov 向前方程积分，就得到 Master 方程.

在统计物理中，对于概率性态的发展，常常需要求解一个 Master 方程.

注 1　在系数含 t 时，Master 方程对应地显示为

$$\frac{\partial p}{\partial t}=\frac{\partial^2}{\partial y^2}\left(\frac{\sigma(t,y)^2}{2}p\right)-\frac{\partial}{\partial y}(b(t,y)p).$$

注 2　Stratonovich 型随机微分方程

$$\xi_t=x+\int_0^t b(\xi_s)\mathrm{d}s+\int_0^t \sigma(\xi_s)\circ \mathrm{d}B_s$$

的解，可以化归 Ito 型随机微分方程

$$\xi_t=x+\int_0^t\left[b(\xi_s)+\frac{1}{2}\sigma(\xi_s)\sigma'(\xi_s)\right]\mathrm{d}s+\int_0^t\sigma(\xi_s)\mathrm{d}B_s$$

处理，其 Kolmogorov 向前方程(Fokker-Planck 方程)则是如下的散度型的

$$\frac{\partial p}{\partial t}=\frac{\partial}{\partial y}\left(\frac{\sigma(y)^2}{2}\frac{\partial p}{\partial y}\right)-\frac{\partial}{\partial y}\left(\left[b(y)-\frac{1}{2}\sigma(y)\sigma'(y)\right]p\right).$$

警告　有一些书上将它误写为

$$\frac{\partial p}{\partial t}=\frac{\partial}{\partial y}\left(\frac{\sigma(y)^2}{2}\frac{\partial p}{\partial y}\right)-\frac{\partial}{\partial y}(b(y)p),$$

一般地这是不对的，只有当 $\sigma(\cdot)$是常数时才正确.

再则,如果我们定义在划分区间取被积函数的终点值的如下的随机积分:对于 f 为连续可微函数,定义当 $n\to+\infty$ 时

$$\sum_k f(\xi_{t_{k+1}^{(n)}})\Delta B_{t_k^{(n)}}$$

的概率极限,我们称之为"向后随机积分",并将它记为 $\int_0^T f(B_t)*\mathrm{d}B_t$. 那么,对于由这种向后随机积分定义的随机微分方程

$$\xi_t=x+\int_0^t b(\xi_s)\mathrm{d}s+\int_0^t\sigma(\xi_s)*\mathrm{d}B_s,$$

可以类似地得到一个扩散过程,其散度型的 Kolmogorov 向前方程(Fokker-Planck 方程)正是

$$\frac{\partial p}{\partial t}=\frac{\partial}{\partial y}\left(\frac{\sigma(y)^2}{2}\frac{\partial p}{\partial y}\right)-\frac{\partial}{\partial y}(b(y)p).$$

一般地说,散度型的优点是,其二阶导数部分写为一个形式对称的运算,在二阶线性抛物型偏微分方程中,这种形式是更为容易处理.

总之,对于随机积分有三种不同定义,它们各有优缺点:Ito 积分(向前随机积分)的优点是具有鞅性质,因而有利于处理轨道性质和平均性质. Stratonovich 积分(中间点随机积分)的优点是有与通常微积分一样形式的复合函数的链法则和用光滑函数近似白噪声时的稳定性. 而向后随机积分的优点是 Fokker-Planck 方程的散度型具有简单而直接的形式.

7.3 多维扩散过程

7.3.1 多维扩散方程

以上的有关扩散过程的结论,都可以推广到多维情形.

定理 7.7 d 维随机微分方程的解,仍然是 Markov 过程,称为 ***d* 维扩散过程**,其转移密度存在,而且满足以下诸方程:

Kolmogorov 向后方程

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum_{i,j=1}^d a_{ij}(\boldsymbol{x})\frac{\partial^2 p}{\partial x_i\partial x_j}+\boldsymbol{b}(\boldsymbol{x})\cdot\nabla_x p,$$

其中

$$\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d,\quad (a_{ij}(\boldsymbol{x}))_{i,j\leqslant d}=\boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}},$$

∇_x是关于变量 $\boldsymbol{x}$ 的梯度运算.

Kolmogorov 向前方程

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum_{i,j=1}^d\frac{\partial^2}{\partial y_i\partial y_j}(a_{ij}(\boldsymbol{y})p)-\nabla\cdot(\boldsymbol{b}(\boldsymbol{y})p),$$

其中$\nabla\cdot$是对于变量 $\boldsymbol{y}$ 的散度运算.

扩散过程$\boldsymbol{\xi}_t$ 在时刻 t 的分布密度 $p(t,\boldsymbol{x})$还满足 Master 方程

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial y_i\partial y_j}(a_{ij}(\boldsymbol{y})p)-\nabla\cdot(\boldsymbol{b}(\boldsymbol{y})p).$$

注　对于 Stratonovich 型随机微分方程

$$\boldsymbol{\xi}_t=\boldsymbol{x}+\int_0^t\boldsymbol{b}(\boldsymbol{\xi}_s)\mathrm{d}s+\int_0^t\boldsymbol{\Sigma}(\boldsymbol{\xi}_s)\circ\mathrm{d}\boldsymbol{B}_s,$$

其 Kolmogorov 向后方程为

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial y_i\partial y_j}(a_{ij}(\boldsymbol{x})p)+\sum_{j=1}^{d}\left(b_j(\boldsymbol{y})+\frac{1}{4}\sum_{i=1}^{d}\frac{\partial a_{ij}}{\partial y_i}\right)\frac{\partial p}{\partial y_j}.$$

而 Kolmogorov 向前方程可以写为以下的**散度形式**：

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial}{\partial y_i}\left(a_{ij}(\boldsymbol{y})\frac{\partial p}{\partial y_j}\right)-\sum_{j=1}^{d}\frac{\partial}{\partial y_j}\left[\left(b_j(\boldsymbol{y})-\frac{1}{4}\sum_{i=1}^{d}\frac{\partial a_{ij}}{\partial y_i}\right)p\right].$$

7.3.2　非退化时齐扩散过程的两歧性和不变密度

在本段中,我们讨论系数不含 t 的时齐情形.

定义 7.2　如果存在非负函数 $\varphi(x)$,使时齐的 Markov 过程的转移密度 $p(t,x,y)$满足

$$\varphi(y)=\int p(t,x,y)\varphi(x)\mathrm{d}x,$$

那么,我们就将 $\varphi(x)$称为转移密度 $p(t,x,y)$的**不变密度**,也称为 Markov 过程的不变密度.

例 7.5　$\varphi(x)\overset{\text{def}}{=\!=}\dfrac{1}{\sigma\sqrt{\dfrac{\pi}{b}}}\mathrm{e}^{-\frac{x^2}{\frac{\sigma^2}{b}}}$是 OU 过程的不变密度.

证明　注意

$$\lim_{t\to+\infty}p(t,x,y)=\frac{1}{\sigma\sqrt{\dfrac{\pi}{b}}}\mathrm{e}^{-\frac{y^2}{\frac{\sigma^2}{b}}}.$$

在 Chapman-Kolmogorov 方程中令 $s\to+\infty$,并将变量 z 改记为 x,便得到

$$\varphi(y)=\int p(t,x,y)\varphi(x)\mathrm{d}x.$$

类似地我们可以得到：如果当 $t\to+\infty$时 Markov 过程的转移密度有不依赖初始状态 x 的非零极限,记 $\varphi(y)=\lim\limits_{t\to+\infty}p(t,x,y)$,那么 $\varphi(x)$就是转移密度的不变密度.

与之相反,Brown 运动却不存在不变密度,因为其转移密度

$$b(t,x,y)=\frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{(y-x)^2}{2t}}\xrightarrow{t\to+\infty}0.$$

然而,对于 $\sigma(x)$ 有正下界(一般称为**非退化的**)的扩散过程,有一个公认的论断,就是对于较好的系数 $\sigma(x)$, $b(x)$,作为随机微分方程解的扩散过程的转移密度具有如下的 Orey 的**两歧性质**:在 $t\to+\infty$ 时,要么 $p(t,x,y)\to 0$(例如 Brown 运动),要么 $p(t,x,y)\to$ 某个与 x 无关 $\varphi(y)$,且满足 $\int\varphi(y)\mathrm{d}y=1$. 这两种情形分别类似于 Markov 链的非常返、零常返($p(t,x,y)\to 0$)与正常返($p(t,x,y)\to\varphi(y)$). 然而这个断言完整而严格的数学证明却并没有完全流行.

在第二种情形,根据偏微分方程的理论,极限函数 φ 应该是 Kolmogorov 向前方程的定态解(即不依赖时间 t 的解),即满足椭圆型方程

$$\frac{\partial^2}{\partial y^2}\left(\frac{\sigma(y)^2}{2}\varphi\right)-\frac{\partial}{\partial y}(b(y)\varphi)=0.$$

另一方面,在 Chapman-Kolmogorov 方程 $p(t+s,x,y)=\int p(t,x,z)p(s,z,y)\mathrm{d}z$ 中,令 $t\to+\infty$,便得 $\varphi(y)=\int\varphi(z)p(s,z,y)\mathrm{d}z$. 这说明了极限函数 φ 是扩散过程的转移函数的不变密度. 直观地看,多维扩散过程的转移密度的不变密度 φ 应该是多维的 Master 方程的稳态解. 可以由严格地数学证明,在系数很宽松的要求下,这确实是对的. 也就是不变密度 φ 满足方程:

$$\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial y_i\partial y_j}(a_{ij}(\boldsymbol{y})\varphi)-\nabla\cdot(\boldsymbol{b}(\boldsymbol{y})\varphi)=0.$$

综上就归结出扩散过程的以下的重要结论.

定理 7.8(转移密度的两歧性) 通常的(例如,系数光滑,扩散系数有一个正的下界的情形)随机微分方程,其解 d 维扩散过程 $\{\xi_t,t\geqslant 0\}$ 的转移密度 $p(t,x,y)$ 满足:或者 $\lim\limits_{t\to+\infty}p(t,x,y)=$ 某个不变密度 $\varphi(x)$,或者 $\lim\limits_{t\to+\infty}p(t,x,y)=0$. 而且有不变密度的充要条件为方程

$$\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial y_i\partial y_j}(a_{ij}(\boldsymbol{y})\varphi)-\nabla\cdot(\boldsymbol{b}(\boldsymbol{y})\varphi)=0$$

有非负的非零可积解. 在条件成立时,$\varphi\xlongequal{\text{def}}\dfrac{\psi}{\int\psi}$ 就是扩散过程的不变密度. 这时,如果初值 ξ_0 的密度是 $\varphi(x)$,那么 $\{\xi_t,t\geqslant 0\}$ 是平稳过程.

证明 最后的结论证明如下:由于 $p(t,x,y)$ 是不变密度,所以,对于任意 t,随机变量 ξ_t 的概率密度也是 $\varphi(x)$. 于是利用 Markov 过程的时齐性,就得到 $(\xi_{t_1+s},\xi_{t_2+s},\cdots,\xi_{t_n+s})$ 在点 $(x_1,x_2,\cdots,x_n)$ 的联合密度为 $\pi(x_1)p(t_2-t_1,x_1,x_2)\cdots p(t_n-t_{n-1},x_{n-1},x_n)$. 可见它与 s 无关. 从而它与 $(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$ 有相同的联合密度.

7.4　扩散过程的遍历定理

由于扩散过程的遍历定理在物理、化学、生物、工程、经济、金融等诸多领域有广泛的应用，我们不加证明地叙述以下两个定理.

定理 7.9(个体遍历定理)　如果存在 $0<\gamma_1<\gamma_2$，使扩散矩阵 $\boldsymbol{A}(\boldsymbol{x})=(a_{ij}(\boldsymbol{x}))$满足

$$\gamma_1 \boldsymbol{I} \leqslant \boldsymbol{A}(\boldsymbol{x}) \leqslant \gamma_2(1+|\boldsymbol{x}|^2)\boldsymbol{I} \quad (\boldsymbol{I}\text{ 是单位矩阵}),$$

而且方程

$$\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial y_i \partial y_j}(a_{ij}(\boldsymbol{y})\varphi)-\nabla\cdot(\boldsymbol{b}(\boldsymbol{y})\varphi)=0$$

有一个非负非零的可积解. 那么其归一化 $\varphi(\boldsymbol{x})$是对应的扩散过程 ξ_t 的不变密度，并且成立如下的遍历定理：对于任意有界的，或非负而 $\int f(x)\varphi(x)\mathrm{d}x<+\infty$ 的(可以无界) Borel 函数 f，不管 ξ_0 的初始分布是什么，恒有

$$P\left(\frac{1}{T}\int_0^T f(\xi_t)\mathrm{d}t \xrightarrow{T\to+\infty} \int f(x)\varphi(x)\mathrm{d}x\right)=1.$$

这个定理的证明远远超出本书的水平. 故而从略. 此定理是 Markov 链的遍历定理的扩散过程版本，其出处可参见钱敏平，龚光鲁所编著的《随机过程论》(第二版，北京大学出版社，1997). 该书中叙述的遍历性条件，是由 Orey 的二歧性定理保证成立的.

类似的事实对于多个时刻也成立，即对于任意 n，任意 $0<s_1<\cdots<s_n$ 及任意的有界 Borel 函数 $f(x_1,x_2,\cdots,x_n)$，不管 ξ_0 的初始分布是什么，总以概率为 1 地有

$$\begin{aligned}&\lim_{T\to+\infty}\frac{1}{T}\int_0^T f(\xi_{t+s_1},\xi_{t+s_2},\cdots,\xi_{t+s_n})\mathrm{d}t\\ =&\int f(x_1,x_2,\cdots,x_n)\varphi(x)p(s_2-s_1,x_1,x_2)\cdots p(s_n-s_{n-1},x_{n-1},x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n.\end{aligned}$$

我们强调这里的遍历性定理的结论不依赖过程的初始状态，这正是统计物理所需要的“各态历经性质”. 正是利用了 Markov 性的特点而得到的. 对比起来，即使用强有力的平稳过程的遍历理论也不能得到这个结论.

作为定理 7.9 的推论，我们有下面的定理.

定理 7.10(平均值定理)　在定理 7.9 的条件下，对于任意有界的 Borel 函数 f，及任意初值 $\xi_0=x$ 有

$$\frac{1}{T}\int_0^T E_x f(\xi_t)\mathrm{d}t \xrightarrow{T\to+\infty} \int f(y)\varphi(y)\mathrm{d}y.$$

证明从略.

7.5　多维扩散过程的首达时与首达地点的分布

我们限于考虑系数不含时间 t 的情形. 对于 d 维扩散过程,如下确定的(向后)微分算子 L

$$Lu = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\boldsymbol{x})\frac{\partial^2 u}{\partial x_i \partial x_j} + \boldsymbol{b}(\boldsymbol{x}) \cdot \nabla u$$

十分重要,它描述了过程的很多分析特性,其中

$$\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d, \quad (a_{ij}(\boldsymbol{x}))_{i,j\leqslant d} = \boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}}, \quad \nabla\text{是梯度运算}.$$

令 Ω 是$\mathbb{R}^d$ 中的一个有界开区域,其边界记为$\partial\Omega$.

7.5.1　多维扩散过程的首达时与吸收过程

以 τ_x 记初值$\boldsymbol{\xi}_0=\boldsymbol{x}(\boldsymbol{x}\in\Omega)$时的扩散过程首达边界$\partial\Omega$ 的时间,即

$$\tau_{\boldsymbol{x}} = \begin{cases} \inf\{t: \boldsymbol{\xi}_t \in \partial\Omega\}, & \text{如果}\{t: \boldsymbol{\xi}_t \in \partial\Omega\}\text{ 不是空集}, \\ +\infty, & \text{如果}\{t: \boldsymbol{\xi}_t \in \partial\Omega\}\text{ 是空集}. \end{cases}$$

时齐的扩散过程$\boldsymbol{\xi}_t$ 在到达边界前,即 $t<\tau$ 的部分为

$$\boldsymbol{\xi}_t^{\partial\Omega} \overset{\text{def}}{=\!=} \boldsymbol{\xi}_t \quad (t<\tau)$$

称为**在边界$\partial\Omega$ 吸收的过程**,其中 τ 是扩散过程$\boldsymbol{\xi}_t$ 首次到达边界$\partial\Omega$ 的时刻. 吸收过程的直观含义是:一旦扩散过程到达边界$\partial\Omega$,它就完全消失了(也称为**斩杀过程**). 直观上看这个在边界$\partial\Omega$ 吸收的过程在区域 Ω 内应该与原过程有相同的转移密度,不同之处是,在边界$\partial\Omega$ 吸收的过程$\boldsymbol{\xi}_t^{\partial\Omega}$不再可能到达边界,也就是$\boldsymbol{\xi}_t^{\partial\Omega}$转移到边界$\partial\Omega$ 的概率密度应该为 0. 于是,我们不加证明地得到以下的定理.

定理 7.11　在边界$\partial\Omega$ 吸收的过程$\boldsymbol{\xi}_t^{\partial\Omega}$的转移密度,记为 $p_\Omega(t,\boldsymbol{x},\boldsymbol{y})$,它满足方程

$$\begin{cases} \dfrac{\partial p_\Omega}{\partial t} = L_{\boldsymbol{y}}^* p_\Omega, \\ p_\Omega \mid_{\boldsymbol{y}\in\Omega} = 0, \end{cases}$$

$$\lim_{t\to 0^+}\int p_\Omega(t,\boldsymbol{x},\boldsymbol{y})f(\boldsymbol{y})\mathrm{d}\boldsymbol{y} = f(\boldsymbol{x}) \quad \text{(初值)},$$

其中

$$\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d, \quad (a_{ij}(\boldsymbol{x}))_{i,j\leqslant d} = \boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}},$$

吸收过程的密度函数 $p_\Omega(t,\boldsymbol{x},\boldsymbol{y})$的解析表达式几乎很少可能得到. 但是本段的讨论有两个方面的意义:有助于进行理论方面的分析研究,也有助于通过随机模拟得到过程在区域中的一些特征量的统计估计.

注　在系数含时间 t 时相应地要改为:吸收转移密度 $p_\Omega(s,\boldsymbol{x},t,\boldsymbol{y})$,$s<t$,满足向后

方程

$$\begin{cases}\dfrac{\partial p_\Omega}{\partial s}+L_x p_\Omega=0,\\ p_\Omega\mid_{y\in\Omega}=0,\end{cases}$$

$$\lim_{t\to s^+}\int p_\Omega(s,\boldsymbol{x},t,\boldsymbol{y})f(\boldsymbol{y})\mathrm{d}\boldsymbol{y}=f(\boldsymbol{x})\quad(\text{初值}).$$

问题1　求 $E\tau_{\boldsymbol{x}}$.

定理7.12　若以下的 **Dynkin** 方程

$$\begin{cases}Lu=-1, & \boldsymbol{x}\in\Omega,\\ u=0, & \boldsymbol{x}\in\partial\Omega\end{cases}$$

有一个二次连续可微的有界正解 $v(\boldsymbol{x})$，则到达边界的平均时间 $E\tau_{\boldsymbol{x}}=v(\boldsymbol{x})$.

证明　对于 $v(\boldsymbol{\xi}_t)$用 Ito 公式，有

$$v(\boldsymbol{\xi}_t)=v(\boldsymbol{x})+\int_0^t Lv(\boldsymbol{\xi}_s)\mathrm{d}s+\int_0^t\sigma^{\mathrm{T}}(\boldsymbol{\xi}_s)\cdot\nabla v(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s.$$

取 $t=\tau_x$ 得到

$$\begin{aligned}0=v(\boldsymbol{\xi}_{\tau_x})&=v(\boldsymbol{x})+\int_0^{\tau_x}Lv(\boldsymbol{\xi}_s)\mathrm{d}s+\int_0^{\tau_x}\sigma^{\mathrm{T}}(\boldsymbol{\xi}_s)\cdot\nabla v(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s\\ &=v(\boldsymbol{x})-\tau_x+\int_0^{\tau_x}\sigma^{\mathrm{T}}(\boldsymbol{\xi}_s)\cdot\nabla v(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s\end{aligned}$$

取期望得

$$0=v(\boldsymbol{x})-E\tau_{\boldsymbol{x}}.$$

即

$$E\tau_{\boldsymbol{x}}=v(\boldsymbol{x}).$$

利用这个公式可以通过初值为 $\boldsymbol{x}$ 过程的多个独立样本$\{\boldsymbol{\xi}_t^{(k)},t\geqslant 0,\xi_0^{(k)}=\boldsymbol{x}\},k\leqslant n$，由$\dfrac{\tau_x^{(1)}+\tau_x^{(2)}+\cdots+\tau_x^{(n)}}{n}$模拟 Dynkin 方程的解 $v(\boldsymbol{x})$.

注　在系数含时间 t 时相应地要改为：如果 Dynkin 方程

$$\begin{cases}\dfrac{\partial u}{\partial t}+Lu=-1, & \boldsymbol{x}\in\Omega,\\ u(t,\boldsymbol{x})=0, & \boldsymbol{x}\in\partial\Omega,t\geqslant s\end{cases}$$

有一个对于 t 一次连续可微，对于 $\boldsymbol{x}$ 二次连续可微的有界正解 $v(t,\boldsymbol{x})$，则 $E\tau_{s,\boldsymbol{x}}=s+v(s,\boldsymbol{x})$.

问题2　求 τ_x 的分布.

定理7.13　如果方程

$$\begin{cases}\dfrac{\partial u}{\partial t}+Lu=0, & \boldsymbol{x}\in\Omega,t\leqslant T,\\ u(T,\boldsymbol{x})=0, & \boldsymbol{x}\in\Omega,\\ u(t,\boldsymbol{x})=1. & \boldsymbol{x}\in\partial\Omega,t<T\end{cases}$$

有对 t 一次连续可微,对 $\boldsymbol{x}$ 二次连续可微的解.记为 $u_T(t,\boldsymbol{x})$,那么

$$P(\tau_{\boldsymbol{x}} < T) = u_T(0,\boldsymbol{x}).$$

证明 对 $u_T(t,\boldsymbol{\xi}_t)$用 Ito 公式

$$u_T(t,\boldsymbol{\xi}_t) = u_T(0,\boldsymbol{x}) + \int_0^t \nabla u_T(s,\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s + \int_0^t \left(\frac{\partial u_T}{\partial s} + Lu_T\right)(s,\boldsymbol{\xi}_s)\mathrm{d}s.$$

于是有

$$\begin{aligned} u_T(T \wedge \tau_{\boldsymbol{x}},\boldsymbol{\xi}_{t\wedge\tau_{\boldsymbol{x}}}) &= u_T(0,\boldsymbol{x}) + \int_0^{T\wedge\tau_{\boldsymbol{x}}} \nabla u_T(s,\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s \\ &\quad + \int_0^{T\wedge\tau_{\boldsymbol{x}}} \left(\frac{\partial u_T}{\partial s} + Lu_T\right)(s,\boldsymbol{\xi}_s)\mathrm{d}s, \end{aligned}$$

其中 $a\wedge b=\min\{a,b\}$.取期望并利用

$$Eu_T(T \wedge \tau_{\boldsymbol{x}},\boldsymbol{\xi}_{T\wedge\tau_{\boldsymbol{x}}}) = u_T(T,\boldsymbol{x})P(\tau_{\boldsymbol{x}} \geqslant T) + u(\tau_{\boldsymbol{x}},\boldsymbol{\xi}_{\tau_{\boldsymbol{x}}})P(\tau_{\boldsymbol{x}} < T) = P(\tau_{\boldsymbol{x}} < T),$$

得到

$$\begin{aligned} P(\tau_{\boldsymbol{x}} < T) &= Eu_T(T \wedge \tau_{\boldsymbol{x}},\boldsymbol{\xi}_{t\wedge\tau_{\boldsymbol{x}}}) \\ &= u_T(0,\boldsymbol{x}) + E\int_0^{T\wedge\tau_{\boldsymbol{x}}} \left(\frac{\partial u_T}{\partial s} + Lu_T\right)(s,\boldsymbol{\xi}_s)\mathrm{d}s = u_T(0,\boldsymbol{x}). \end{aligned}$$

用这个表达式可以通过初值为 $\boldsymbol{x}$ 过程的多个独立样本$\{\boldsymbol{\xi}_t^{(k)},t\geqslant 0,\xi_0^{(k)}=\boldsymbol{x}\},k\leqslant n$,模拟 $\tau_{\boldsymbol{x}}$ 的分布并估计 $u_T(0,\boldsymbol{x})$.

也可以用吸收过程的转移密度表达 $\tau_{\boldsymbol{x}}$ 的分布函数.为此注意吸收过程就是$\boldsymbol{\xi}_t^{\partial\Omega}=\boldsymbol{\xi}_t(t<\tau)$,所以 $P(\tau_{\boldsymbol{x}}>t)$表示在时间 t 前还未被吸收的概率,即

$$P(\tau_{\boldsymbol{x}} > t) = \int p_\Omega(t,\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{y}.$$

故而 $\tau_{\boldsymbol{x}}$ 的分布函数是

$$P(\tau_{\boldsymbol{x}} \leqslant t) = 1 - \int p_\Omega(t,\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{y}.$$

而且我们还可以得到 $E\tau_{\boldsymbol{x}}$ 的另一个表达式:

$$E\tau_{\boldsymbol{x}} = \int_0^{+\infty} P(\tau_{\boldsymbol{x}} > t)\mathrm{d}t = \int_0^{+\infty}\int p_\Omega(t,\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{y}\mathrm{d}t.$$

注 在系数含时间 t 时相应地改为:如果方程

$$\begin{cases} \dfrac{\partial u}{\partial t} + Lu = 0, & x \in \Omega, t \leqslant T, \\ u(T,x) = 0, & x \in \Omega, \\ u(t,x) = 1, & x \in \partial\Omega, s \leqslant t \leqslant T \end{cases}$$

有对 t 一次连续可微,对 x 二次连续可微的解.记为 $u_T(t,\boldsymbol{x})$,那么

$$P(\tau_{\boldsymbol{x}} < T) = u_T(s,\boldsymbol{x}).$$

7.5.2 多维扩散过程的首达地点分布与 Dirichlet 边值问题

问题 1 Dirichlet 边值问题解的概率表示.

定理 7.14 若椭圆型方程的 Dirichlet 边值问题

$$\begin{cases} Lu = 0, & \boldsymbol{x} \in \Omega, \\ u(\boldsymbol{x}) = f(\boldsymbol{x}). & \boldsymbol{x} \in \partial\Omega \end{cases}$$

在 Ω 有二次连续可微的解 $u(\boldsymbol{x})$,则有 Kolmogorov 公式

$$u(\boldsymbol{x}) = Ef(\boldsymbol{\xi}_{\tau_x}).$$

这是调和函数平均值公式的概率形式推广.

证明 注意$\boldsymbol{\xi}_0 = \boldsymbol{x}$,对 $u(\boldsymbol{\xi}_t)$用 Ito 公式

$$u(\boldsymbol{\xi}_t) = u(\boldsymbol{x}) + \int_0^t \nabla u(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s + \int_0^t (Lu)(\boldsymbol{\xi}_s)\mathrm{d}s.$$

于是有

$$f(\boldsymbol{\xi}_{\tau_x}) = u(\boldsymbol{\xi}_{\tau_x}) = u(\boldsymbol{x}) + \int_0^{\tau_x} \nabla u(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s + \int_0^{\tau_x} Lu(\boldsymbol{\xi}_s)\mathrm{d}s = u(\boldsymbol{x}) + \int_0^{\tau_x} \nabla u(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s.$$

取期望得到

$$Ef(\boldsymbol{\xi}_{\tau_x}) = u(\boldsymbol{x}).$$

利用这个公式,可以通过随机模拟得到有界区域的 Dirichlet 问题解在指定点 x 的值的估计:首先生成过程$\{\boldsymbol{\xi}_t, t\geqslant 0\}$的 n 个模拟样本$\{\boldsymbol{\xi}_t^{(i)}, t\geqslant 0\}, i\leqslant n$. 那么,以很大的概率有

$$u(\boldsymbol{x}) \approx \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{\xi}_{\tau_x}^{(i)}).$$

当然,在实际施行时还要将空间离散化.

问题 2 首达地点的分布.

定理 7.15 对于边界$\partial\Omega$ 上的一个部分 Γ,如果方程

$$\begin{cases} Lu = 0, & \boldsymbol{x} \in \Omega, \\ u(x) = 1, & \boldsymbol{x} \in \Gamma, \\ u(x) = 0, & \boldsymbol{x} \in \partial\Omega - \Gamma \end{cases}$$

在 Ω 有二次连续可微的解 $u_\Gamma(\boldsymbol{x})$,则有

$$P_x(\boldsymbol{\xi}_{\tau_x} \in \Gamma) = u(\boldsymbol{x}).$$

证明 这是问题 1 的特殊情形,即 $f = I_\Gamma$ 的情形.

*7.6 Girsanov 定理与 Feyman-Kac 公式

Girsanov 定理与 Feyman-Kac 公式是随机微积分中,除了 Ito 公式以外两个最重要的基本定理. 它们与 Ito 公式一起,形成了随机分析理论与应用的三大支柱. 它们的证明

需要随机分析的工具，所以我们只给出叙述，而不给证明.

7.6.1 Brown 的随机平移——Girsanov 变换

定义 7.3(Girsanov 变换) 设 $\boldsymbol{B}_t$ 为概率空间$(\Omega,\mathcal{F},P)$上的 d 维 Brown 运动，$\boldsymbol{\Phi}_t$ 为$(\boldsymbol{B}_t)$可知的有界随机过程. 令

$$\hat{\boldsymbol{B}}_t = \boldsymbol{B}_t - \int_0^t \boldsymbol{\Phi}_s \mathrm{d}s,$$

它称为 **Brown 运动的随机平移或 Girsanov 变换**. 定义一个新的概率 P^* 如下：对于$\mathcal{F}$中的任意事件 A，只要它的信息完全可由$\{B_s: s\leqslant t\}$确定，就定义它的新的概率为

$$P^*(A) = E[I_A \mathrm{e}^{\int_0^t \boldsymbol{\Phi}_s^{\mathrm{T}} \mathrm{d}\boldsymbol{B}_s - \frac{1}{2}\int_0^t \boldsymbol{\Phi}_s^{\mathrm{T}} \boldsymbol{\Phi}_s \mathrm{d}s}],$$

则不难验证 P^* 满足概率的公理，因而它确是$\mathcal{F}$上的一个概率. 下面是一个极为深刻的定理.

定理 7.16(**Girsanov** 定理) 在新概率 P^* 下，Brown 运动的 Girsanov 变换$\hat{\boldsymbol{B}}_t$ 是 d 维 Brown 运动(警告：在原来的概率 P 下它并不是 Brown 运动).

Girsanov 定理有广泛的应用. 例如，把它应用到金融数学中的 Black-Scholes 模型，就可以得到风险中性的概率，这是期权(option)定价的理论基础.

7.6.2 Feyman-Kac 公式

定理 7.17(Feyman-Kac 公式) 设$\boldsymbol{\xi}_t$ 是随机微分方程 $\boldsymbol{\xi}_t = \boldsymbol{x} + \int_0^t \boldsymbol{b}(\boldsymbol{\xi}_s)\mathrm{d}s + \int_0^t \boldsymbol{\Sigma}(\boldsymbol{\xi}_s)\mathrm{d}\boldsymbol{B}_s$ 的解，$c(x)\geqslant 0$，$(a_{ij}(x))_{i,j\leqslant d} = \boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}}$. 那么，在 $0\leqslant t\leqslant T$ 满足终值条件

$$u(T,\boldsymbol{x}) = f(\boldsymbol{x}) \quad (\boldsymbol{x}\in\mathbb{R}^d)$$

的线性偏微分方程

$$\frac{\partial u}{\partial t} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(\boldsymbol{x})\frac{\partial^2 u}{\partial x_i \partial x_j} + \boldsymbol{b}(\boldsymbol{x})\nabla_x u - c(\boldsymbol{x})u + g(t,\boldsymbol{x})$$

的解可以用概率表示为以下的期望的形式：

$$u(t,\boldsymbol{x}) = E\left(\int_t^T \mathrm{e}^{-\int_t^s c(\boldsymbol{\xi}_u)\mathrm{d}u} g(s,\boldsymbol{\xi}_s)\mathrm{d}s + \mathrm{e}^{-\int_t^T c(\boldsymbol{\xi}_u)\mathrm{d}u} f(\boldsymbol{\xi}_T)\right).$$

Feyman-Kac 公式将一个线性偏微分方程的终值问题的解，表达为一个随机微分方程的解的函数的条件期望. 于是可以通过随机模拟得到线性偏微分方程的终值问题的数值解. 例如说，得到

$$u(0,\boldsymbol{x}) = E\left[\int_t^T \mathrm{e}^{-\int_t^s c(\boldsymbol{\xi}_u)\mathrm{d}u} g(s,\boldsymbol{\xi}_s)\mathrm{d}s + \mathrm{e}^{-\int_t^T c(\boldsymbol{\xi}_u)\mathrm{d}u} f(\boldsymbol{\xi}_T)\right]$$

的近似值.

而 Dirichlet 问题

$$\begin{cases}\dfrac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\boldsymbol{x})\dfrac{\partial^2 u}{\partial x_i \partial x_j}+\boldsymbol{b}(\boldsymbol{x})\nabla_{\boldsymbol{x}} u-c(\boldsymbol{x})u+g(\boldsymbol{x})=0, & \boldsymbol{x}\in G,\\ u(\boldsymbol{x})=0, & \boldsymbol{x}\in\partial G\end{cases}$$

的解的概率表示则为

$$u(\boldsymbol{x})=E\left(\int_0^{\tau_x} g(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\mathrm{d}t\,\middle|\,\boldsymbol{\xi}_0=\boldsymbol{x}\right),$$

其中 τ_x 是随机过程 $\boldsymbol{\xi}_t$ 首达边界 ∂G 的时刻.

我们对 Dirichlet 问题情形给出一个直观说明. 假定 $\boldsymbol{\xi}_0=\boldsymbol{x}$. 由 Ito 公式得到

$$\begin{aligned}\mathrm{d}\left(u(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\right)&=\mathrm{d}u(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}+u(\boldsymbol{\xi}_t)\mathrm{d}\left(\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\right)\\&=\left[\left(\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\boldsymbol{x})\frac{\partial^2 u}{\partial x_i \partial x_j}+\boldsymbol{b}(\boldsymbol{x})\nabla_{\boldsymbol{x}} u\right)\mathrm{d}t+\boldsymbol{b}(\boldsymbol{x})\nabla_{\boldsymbol{x}} u\,\mathrm{d}\boldsymbol{B}_t\right]_{\boldsymbol{x}=\boldsymbol{\xi}_t}\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\\&\quad+u(\boldsymbol{\xi}_t)(-c(\boldsymbol{\xi}_t)\mathrm{d}t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\end{aligned}$$

其积分形式为

$$\begin{aligned}&u(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}-u(\boldsymbol{x})\\&=\int_0^t\left[\left(\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\boldsymbol{x})\frac{\partial^2 u}{\partial x_i \partial x_j}+\boldsymbol{b}(\boldsymbol{x})\nabla_{\boldsymbol{x}} u\right]_{\boldsymbol{x}=\boldsymbol{\xi}_t}-c(\boldsymbol{\xi}_t)u(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\right)\mathrm{d}t\\&\quad+\int_0^t[\boldsymbol{b}(\boldsymbol{x})\nabla_{\boldsymbol{x}} u]_{\boldsymbol{x}=\boldsymbol{\xi}_t}\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\mathrm{d}\boldsymbol{B}_t.\end{aligned}$$

令 $t=\tau_x$，再取期望，形式地认为 Ito 积分部分的期望为 0，并注意 $\boldsymbol{\xi}_{\tau_x}\in\partial G$，$u(\boldsymbol{\xi}_{\tau_x})=0$，并且 $u(\boldsymbol{x})$ 是 Dirichlet 问题的解，我们便得到

$$u(\boldsymbol{x})=E\left[\int_0^{\tau_x} g(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\mathrm{d}t\right].$$

Feyman-Kac 公式的概率含义

假设一个按随机微分方程的解的轨道作随机运动的粒子在一个小的时间区间 $[t,t+h)$ 以概率 $c(\boldsymbol{\xi}_t)h+o(h)$ 被“斩杀”. 对时间取等分

$$0=t_0<t_1<\cdots<t_n=t,\quad h=t_{k+1}-t_k=\frac{t}{n},$$

那么直到时间 t 为止还存活的概率近似地是

$$\prod_{k=0}^{n-1}(1-c(\boldsymbol{\xi}_{t_k})h)=\exp\left\{\sum_{k=1}^{n-1}\log(1-c(\boldsymbol{\xi}_{t_k})h)\right\}$$

$$\approx \exp\left\{-\sum_{k=1}^{n-1} c(\boldsymbol{\xi}_{t_k})h\right\} \xrightarrow{n\to+\infty} \exp\left\{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s\right\}.$$

于是

$$u(\boldsymbol{x}) = E\left[\int_0^{\tau_x} g(\boldsymbol{\xi}_t)\mathrm{e}^{-\int_0^t c(\boldsymbol{\xi}_s)\mathrm{d}s}\mathrm{d}t\right]$$

= 沿随机轨道在到达边界 ∂G 前还未被斩杀的粒子的轨道函数 $g(\boldsymbol{\xi}_t)$ 值的平均.

*7.7 扩散过程的最佳停止

随机微分方程的解的最佳停止问题,来自最佳控制、最佳估计、美式期权的定价等实际问题.假定一个系统用随机微分方程建模,而这个系统中还含有一个或多个可以调节的参数,称为控制变量,调节这些参数以控制这个系统的动态发展(以达到预先设计的目标),或者只决定是否停止.最佳控制问题就是要使所付出的价格最小.

最简单的情形是只决定是否停止.

考虑 d 维具有光滑边界 ∂G 的开集 G.假定随机微分方程

$$\mathrm{d}\boldsymbol{\xi}_t = \boldsymbol{b}(t,\boldsymbol{\xi}_t)\mathrm{d}t + \boldsymbol{\Sigma}(t,\boldsymbol{\xi}_t)\mathrm{d}\boldsymbol{B}_t$$

对于给定的初值 $\boldsymbol{\xi}_0\in G$ 存在唯一的解. $\boldsymbol{\xi}_t$ 首达边界 ∂G 的时刻记为 τ_x.假定将过程停止在某个停时 θ 需要付出价格: $\int_0^{\theta\wedge\tau_x} f(\boldsymbol{\xi}_t)\mathrm{d}t + g(\boldsymbol{\xi}_\tau)$,其中前一项是运行价格,而后一项是停止的费用.将停止在停时 θ 需要付出的平均价格记为 $J_{\boldsymbol{x}}(\theta)$.于是

$$J_{\boldsymbol{x}}(\theta) = E\left(\left[\int_0^{\theta\wedge\tau_x} f(\boldsymbol{\xi}_t)\mathrm{d}t + g(\boldsymbol{\xi}_{\theta\wedge\tau_x})\right]\,\middle|\,\boldsymbol{\xi}_0 = \boldsymbol{x}\right).$$

最佳停止问题就是要找停时 θ^* 使

$$J_{\boldsymbol{x}}(\theta^*) = \inf_{\theta\text{是停时}} J_{\boldsymbol{x}}(\theta)$$

右边的这个量,称为价格函数,记为 $u(\boldsymbol{x})$,即

$$u(\boldsymbol{x}) \overset{\text{def}}{=\!=} \inf_{\theta\text{是停时}} J_{\boldsymbol{x}}(\theta), \quad \boldsymbol{x}\in G.$$

它是给定起点 $\boldsymbol{x}$ 时的最小期望价格.对于已知的价格函数找最佳停时 θ^* 的常见的方法是动态规划方法.假定 $u(\boldsymbol{x})$十分光滑,我们要找出它满足的条件.首先,有 $u(\boldsymbol{x})\leqslant J_{\boldsymbol{x}}(0)=g(\boldsymbol{x})$,即

(1) $u(\boldsymbol{x})\leqslant g(\boldsymbol{x}),\boldsymbol{x}\in G$.

再则,如果 $\boldsymbol{x}\in\partial G$,那么 $\tau_{\boldsymbol{x}}=0$,所以有 $u(\boldsymbol{x})=g(\boldsymbol{x})$,即

(2) $u(\boldsymbol{x})=g(\boldsymbol{x}),\boldsymbol{x}\in\partial G$.

对于任意 $\boldsymbol{x}\in G$ 及任意一个小的停时 δ,假定在时刻 δ 还没有达到 ∂G,如果我们不在时刻 δ 停止,那么可以考虑这个随机微分方程从新的状态 $\boldsymbol{\xi}_\delta$ 出发,这时的价格函数将是 $u(\boldsymbol{\xi}_\delta)$,于是价格至少有

$$E\left(\int_0^\delta f(\boldsymbol{\xi}_t)\mathrm{d}t + u(\boldsymbol{\xi}_\delta)\right),$$

所以

$$u(\boldsymbol{x}) \leqslant E\left(\int_0^\delta f(\boldsymbol{\xi}_t)\mathrm{d}t + u(\boldsymbol{\xi}_\delta)\right).$$

故而

$$\lim_{\delta\to 0}\frac{1}{\delta}E\left(\int_0^\delta f(\boldsymbol{\xi}_t)\mathrm{d}t + [u(\boldsymbol{\xi}_\delta) - u(\boldsymbol{x})]\right) \geqslant 0.$$

用 Ito 公式,得到

$$f(\boldsymbol{x}) + Lu(\boldsymbol{x}) \geqslant 0$$

或者

(3) $-Lu(\boldsymbol{x})\leqslant f(\boldsymbol{x}),\boldsymbol{x}\in G.$

但是,如果在(1)中在 G 中的某些点 $\boldsymbol{x}$ 出现严格的不等式 $u(\boldsymbol{x})<g(\boldsymbol{x})$,那么最佳情形是不要立刻停止(因为立刻停止的价格是 $g(\boldsymbol{x})$,它比应该达到的最小值 $u(\boldsymbol{x})$大). 直观地设想我们需要让过程进行某个至少很小的一段时间,并使

$$u(\boldsymbol{x}) = E\left(\int_0^\delta f(\boldsymbol{\xi}_t)\mathrm{d}t + u(\boldsymbol{\xi}_\delta)\right),$$

也就是有

$$-Lu(\boldsymbol{x}) = f(\boldsymbol{x}), \quad \text{在满足 } u(\boldsymbol{x}) < g(\boldsymbol{x}) \text{ 的 } \boldsymbol{x} \text{ 点}.$$

综上,我们得到价格函数 $u(\boldsymbol{x})$满足的条件

$$\begin{cases}\max\{-Lu - f, u - g\} = 0, & \boldsymbol{x} \in G,\\ u = g, & \boldsymbol{x} \in \partial G\end{cases}$$

称为最佳条件.

价格函数的求解

我们不加证明地引述下面两个定理.

定理 7.18 假定 f,g 都是光滑函数,那么存在唯一的具有二阶有界导数(但是未必连续)的函数 u 满足:

(1) $u(\boldsymbol{x})\leqslant g(\boldsymbol{x}),\boldsymbol{x}\in G$;

(2) $u(\boldsymbol{x})=g(\boldsymbol{x}),\boldsymbol{x}\in \partial G$;

(3) $-Lu(\boldsymbol{x})\leqslant f(\boldsymbol{x}),\boldsymbol{x}\in G$;

(4) $\max\{-Lu-f,u-g\}=0$ 对于 $\boldsymbol{x}\in G$ 几乎处处成立.

实际上的近似做法是,引入一个惩罚性的光滑的非负非降凸函数 $\lambda_\varepsilon(\boldsymbol{x})$(假定 $\lambda_\varepsilon(\boldsymbol{x})=0(\boldsymbol{x}\leqslant\boldsymbol{0})$并使$\lim\limits_{\varepsilon\to 0}\lambda_\varepsilon(\boldsymbol{x})=+\infty,(\boldsymbol{x}>\boldsymbol{0})$),考虑对原问题使用惩罚函数得到的近似模型:

$$\begin{cases}-Lu_\varepsilon(\boldsymbol{x}) + \lambda_\varepsilon(\boldsymbol{x})(u_\varepsilon(\boldsymbol{x}) - g(\boldsymbol{x})) = f, & \boldsymbol{x} \in G,\\ u_\varepsilon(\boldsymbol{x}) = g, & \boldsymbol{x} \in \partial G.\end{cases}$$

这个近似模型将具有活动边界的微分不等式近似地用一个普通边界条件的偏微分方程. 对于给定的 $\lambda_\varepsilon(\boldsymbol{x})$,可以用数值近似计算上述方程的解.

最佳停止策略的获得

一旦得到了价格函数,我们定义如下的停止集合:

$$S \stackrel{\text{def}}{=\!=} \{\boldsymbol{x} \in G: u(\boldsymbol{x}) = g(\boldsymbol{x})\},$$

它是一个闭集,也就是在 S 中取值的任意序列的极限仍保持在 S 之中.

定理 7.19　假定 f,g 满足定理 7.18 的条件,那么 $u(\boldsymbol{x})$是价格函数,而且对于任意 $\boldsymbol{x}\in\overline{G}=G\cup\partial G$,从 $\boldsymbol{x}$ 出发的随机过程$\boldsymbol{\xi}_t$ 首达 S 的时刻 θ^* 是一个最佳停时,即

$$u(\boldsymbol{x}) = J_{\boldsymbol{x}}(\theta^*) = \inf_{\text{一切停时}\theta} J_{\boldsymbol{x}}(\theta).$$

直观推导　按上面的阐述,在集合 $G-S=\{\boldsymbol{x}\in G: u(\boldsymbol{x})<g(\boldsymbol{x})\}$上应该有 $Lu=f$,而且在 $G-S$ 的边界上有 $u=g$. 注意 $\tau_{\boldsymbol{x}}\wedge\theta^*$ 是随机过程$\boldsymbol{\xi}_t$ 首次越出集合 $G-S$ 的时刻,于是对于 $\boldsymbol{x}\in G-S$ 有

$$u(\boldsymbol{x}) = E\left(\int_0^{\tau_{\boldsymbol{x}}\wedge\theta^*} f(\boldsymbol{\xi}_t)\mathrm{d}t + g(\boldsymbol{\xi}_{\tau_{\boldsymbol{x}}\wedge\theta^*})\right) = J_{\boldsymbol{x}}(\theta^*).$$

但是,对于 $\boldsymbol{x}\in S$ 有 $\tau_{\boldsymbol{x}}\wedge\theta^*\leqslant\theta^*=0$,从而有 $u(\boldsymbol{x})=g(\boldsymbol{x})=J_{\boldsymbol{x}}(\theta^*)$. 综上所述,我们得到:对于任意 $\boldsymbol{x}\in G$ 恒有 $u(\boldsymbol{x})=J_{\boldsymbol{x}}(\theta^*)$.

最后证明 $u(\boldsymbol{x})$确是价格函数. 为此,对于任意停时 θ,在$[0,\tau_x\wedge\theta]$上对 $u(\boldsymbol{\xi}_t)$利用 Ito 公式得到

$$u(\boldsymbol{x}) = E\left[\int_0^{\tau_{\boldsymbol{x}}\wedge\theta} -Lu(\boldsymbol{\xi}_t)\mathrm{d}t + u(\boldsymbol{\xi}_{\tau_{\boldsymbol{x}}\wedge\theta})\right].$$

由于在 $\overline{G}$ 中有 $-Lu\leqslant f, u\leqslant g$,以及上面已经得到的结论,还有

$$J_{\boldsymbol{x}}(\theta^*) = u(\boldsymbol{x}) \leqslant E\left[\int_0^{\tau_{\boldsymbol{x}}\wedge\theta} f(\boldsymbol{\xi}_t)\mathrm{d}t + g(\boldsymbol{\xi}_{\tau_{\boldsymbol{x}}\wedge\theta})\right] = J_{\boldsymbol{x}}(\theta).$$

由停时 θ 的任意性立刻得到 $J_{\boldsymbol{x}}(\theta^*)=u(\boldsymbol{x})=\inf\limits_{\text{停时}\theta} J_{\boldsymbol{x}}(\theta)$.

习题 7

1. 对于有界的实值函数 $f(t)$, 求 $E\mathrm{e}^{\int_0^t f(s)\mathrm{d}B_s}$.

2. 假定 $u(t,x)$满足向后方程$\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial^2 u}{\partial x^2}=0$,证明 $Eu(t,B_t)=u(0,0)$.

3. 求一维初值问题

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},$$

$$u(0,x) = f(x)$$

的解 $f(t,x)$. 再推广至多维情形.

4. 求一维终值问题

$$\begin{cases}\dfrac{\partial u}{\partial t}+\dfrac{\partial^2 u}{\partial x^2}=0,\\ u(T,x)=f(x)\end{cases}$$

的解 $f(t,x)$, $0\leqslant t\leqslant T$. 再推广至多维情形.

第 8 章　随机微分方程的解的数值模拟算法

8.1　随机微分方程轨道的采样近似

随机微分方程描述了相当广泛的一类常见的随机系统. 给出了随机系统建模的一种重要的途径. 统计物理中通过求解 Master 方程,或者 Fokker-Plank 方程得到状态的概率密度或状态转移的概率密度,并用后者描述系统的时间发展,应用领域的人们较少地注意到随机过程按样本轨道的发展,甚至不理解研究随机过程按样本轨道发展的必要性. 这样的了解看来有些粗放. 要想了解随机系统的时间发展,更精确的是将状态按时间的发展作为随机试验的基本结果,进行概率和统计研究,这就是随机过程. 事实上,只知道系统在固定时间的条件规律或状态转移的统计规律,常常是不够的,特别地,有时描写转移的统计概率规律的参数通常并不知道其具体的数值,这就需要通过实验的结果统计地推定,即估计. 就是通过用随机过程的一段相当长时间的现实,即轨道的数据来估值. 这种用随机过程的一个轨道表示过程的特征参数的性质,就是用时间平均近似空间平均的性质,统计物理学家通常称之为遍历性质. 统计物理的基础之一就是假定系统的发展具有遍历性质. 然而如果进一步追溯,什么样的模型能具有遍历性质呢? 统计物理中忽视了这一方面的进一步阐述与研究. Markov 过程与随机微分方程的理论完美地回答了这个问题. 这个理论说明在很宽松的条件下,按样本轨道的采样将给出这个随机过程的平稳分布与转移概率等数值的统计估计,同时也能估计系统的特性参数. 这对于认知与分析系统的性能是十分重要的. 而随机微分方程的解的数值模拟正是提供随机过程的近似样本轨道的实际途径. 了解随机过程的精粹正是要了解状态按系统模型所作的时间发展,随机微分方程提供了一个强大的模型与工具.

在原则上,随机微分方程

$$\xi_t = \xi_0 + \int_0^t b(s,\xi_s)\mathrm{d}s + \int_0^t \sigma(s,\xi_s)\mathrm{d}B_s, \tag{8.1}$$

可以用迭代方法给出数值解. 但是,从计算的角度看,这种方法并不经济. 在实践中,我们需要采用更省时间的算法.

由于$\dfrac{\mathrm{d}B_t}{\sqrt{\mathrm{d}t}} \sim N(0,1)$,在许多应用领域中,例如,在金融领域中,人们常常粗略地在随

机微分方程

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dB_t, \quad \xi_0 = x \tag{8.1$'$}$$

中简单地用 $\eta_t\sqrt{dt}$ 代替 dB_t，其中 η_t 是标准正态随机变量，也就是用随机的方程

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)\eta_t\sqrt{dt}, \quad \xi_0 = x \tag{8.2}$$

近似地代替方程(8.1).而求方程(8.2)的样本轨道的数值解，就是给出这个解在离散时间采样点上的值.为此将$[0,T]$作适当小的步长 $\delta=\dfrac{T}{n}$ 的 n 等分：

$$0 = t_0 < t_1 < \cdots < t_n = T,$$

并且用 $b(\xi_{t_k})$，$\sigma(\xi_{t_k})$分别近似地代替 $b(\xi_t)$，$\sigma(\xi_t)$，这样就得到采样值的递推近似：

$$\xi_{t_{k+1}} = \xi_{t_k} + b(\xi_{t_k})\Delta t_k + \sigma(\xi_{t_k})\eta_k\sqrt{\Delta t_k}, \quad \xi_0 = x, \tag{8.3}$$

其中 η_k 是独立同分布的标准正态随机变量.

这样的近似随机采样较为粗放.在用样本轨道的采样给出平稳分布与转移概率的数值的统计估计时，有时需要更精细的近似.

注意，正如第 6 章所指出的那样，对于一个 Ito 随机微分方程，如果将它形式地看成由白噪声驱动的微分方程，并且用光滑的随机过程列近似白噪声$\dfrac{d\boldsymbol{B}_t}{dt}$(例如随机差商的光滑近似)，那么，相应的带有随机参数的常微分方程的解并不趋于这个 Ito 方程的解，而是趋于将 Ito 积分 $\int_0^t \sigma(\xi_t)dB_t$ 用 Stratonovich 积分$\int_0^t \sigma(\xi_t)\circ dB_t$ 代替后所得到的 Stratonovich 随机微分方程的解.这说明用 Ito 随机微分方程建模会失去光滑逼近下的稳定性.另一方面，用 Stratonovich 型随机微分方程建模时，对于白噪声的光滑的随机逼近，其解就有稳定性.所以有人认为从建模的角度，用 Stratonovich 型随机微分方程建模似乎更为合理.

8.2　随机微分方程在固定时刻附近的随机 Taylor 展开与解的差分近似

假定扩散系数 $\sigma(x)$有界且二次连续可微，漂移系数 $b(x)$连续.

我们考虑随机微分方程(8.1)的 ξ_t 在采样时刻 $t_m(m<n)$附近的展开.自然的想法就是对随机的差分进行近似.

在 $t_m<s<t_{m+1}$上对 $\sigma(\xi_s)$用 Ito 公式得到

$$\sigma(\xi_s) - \sigma(\xi_{t_m}) = \int_{t_m}^{s}\left[b(\xi_u)\sigma'(\xi_u) + \frac{1}{2}\sigma^2(\xi_u)\sigma''(\xi'_u)\right]du + \int_{t_m}^{s}\sigma(\xi_u)\sigma'(\xi_u)dB_u. \tag{8.4}$$

注意 $dB_t dt = o(dt)$，故有

$$\begin{aligned}\xi_{t_{m+1}}-\xi_{t_m}&=\int_{t_m}^{t_{m+1}}b(\xi_s)\mathrm{d}s+\int_{t_m}^{t_{m+1}}\sigma(\xi_s)\mathrm{d}B_s\\&=(b(\xi_{t_m})\delta+o(\delta))+\sigma(\xi_{t_m})(B_{t_{m+1}}-B_{t_m})+\int_{t_m}^{t_{m+1}}[\sigma(\xi_s)-\sigma(\xi_{t_m})]\mathrm{d}B_s\end{aligned}\tag{8.5}$$

$$=b(\xi_{t_m})\delta+\sigma(\xi_{t_m})\zeta_m+\int_{t_m}^{t_{m+1}}\int_{t_m}^{s}\sigma(\xi_u)\sigma'(\xi_u)\mathrm{d}B_u\mathrm{d}B_s+o(\delta),\tag{8.6}$$

其中$\{\zeta_m\}$独立同分布且 $\zeta_m=B_{t_{m+1}}-B_{t_m}\sim N(0,\delta)$.

我们首先考虑(8.5)式的近似. 因为

$$E\left[\int_{t_m}^{t_{m+1}}[\sigma(\xi_s)-\sigma(\xi_{t_m})]\mathrm{d}B_s\right]^2\leqslant\max_{t_m<s<t_{m+1}}E[\sigma(\xi_s)-\sigma(\xi_{t_m})]^2\Delta t_m=o(\delta).$$

粗略地(这里不是严格的数学推导)就有

$$\int_{t_m}^{t_{m+1}}[\sigma(\xi_s)-\sigma(\xi_{t_m})]\mathrm{d}B_s=o(\sqrt{\delta}).$$

所以,只要 $\sigma(x)$有界连续(甚至并不需要二次连续可微),我们就由(8.5)式得到随机微分方程精确到 δ 的$\frac{1}{2}$阶(即精度为 $\delta^{\frac{1}{2}}$,称为半阶近似)的如下的差分近似.

8.2.1 半阶差分近似模型(Euler-Maruyama 近似)

令

$$\xi^{(n)}_{t_{m+1}}=\xi^{(n)}_{t_m}+b(\xi^{(n)}_{t_m})\Delta t_m+\sigma(\xi^{(n)}_{t_m})\zeta_m,\quad \xi^{(n)}_{t_0}=x,\tag{8.7}$$

其中$\{\zeta_m\}$独立同分布且 $\zeta_m=\Delta B_{t_m}\sim N(0,\delta)\left(注意\ \Delta t_m=\frac{T}{n}=\delta\right)$. 注意此式恰好就是(8.3)式的递推公式. 但是,与直观地直接得到(8.3)式的递推公式相比,这里的推导方法具有理论方面的长处,即这个方法可以用来延伸到获得更高精度的近似.

较高精度的近似是利用(8.6)式. 注意到

$$\begin{aligned}\int_{t_m}^{t_{m+1}}\int_{t_m}^{s}\sigma(\xi_u)\sigma'(\xi_u)\mathrm{d}B_u\mathrm{d}B_s&=\sigma(\xi_{t_m})\sigma'(\xi_{t_m})\int_{t_m}^{t_{m+1}}\int_{t_m}^{s}\mathrm{d}B_u\mathrm{d}B_s+o(\delta)\\&=\sigma(\xi_{t_m})\sigma'(\xi_{t_m})\int_{t_m}^{t_{m+1}}(B_s-B_{t_m})\mathrm{d}B_s+o(\delta)\\&=\sigma(\xi_{t_m})\sigma'(\xi_{t_m})\left[\frac{1}{2}(B^2_{t_{m+1}}-B^2_{t_m}-\Delta t_m)-B_{t_m}\Delta B_{t_m}\right]+o(\delta)\\&=\frac{1}{2}\sigma(\xi_{t_m})\sigma'(\xi_{t_m})[(\Delta B_{t_m})^2-\Delta t_m]+o(\delta).\end{aligned}\tag{8.8}$$

由(8.6)式和(8.8)式,我们得到随机微分方程精确到 δ 的 1 次方的如下的一阶差分近似.

8.2.2　一阶差分近似模型(Milstein 近似)

令

$$\xi_{t_{m+1}}^{(n)}=\xi_{t_m}^{(n)}+b(\xi_{t_m}^{(n)})\Delta t_m+\sigma(\xi_{t_m}^{(n)})\zeta_m+\frac{1}{2}\sigma(\xi_{t_m}^{(n)})\sigma'(\xi_{t_m}^{(n)})(\zeta_m^2-\Delta t_m),\qquad(8.9)$$

其中$\{\zeta_m\}$独立同分布,$\zeta_m\sim N(0,\delta)$,$X_{t_0}^{(n)}=x$.

具体的做法是:对于给定的 n,通过独立地生成的 $N(0,\delta)$随机数 $\zeta_m(m<n)$,逐步地解出的 $\xi_{t_1}^{(n)},\xi_{t_2}^{(n)},\cdots,\xi_{t_n}^{(n)}$,就得到随机微分方程的近似解在离散时间集合$\{t_1,t_2,\cdots,t_n\}$上的采样值.用它们用折线作内插近似,或用样条函数(spline)作内插近似,就得到原来的随机微分方程的一个近似模拟轨道.

如果漂移系数 $b(x)$与扩散系数有更多阶的连续导数,我们还可以沿着这个思路构造$\frac{3}{2}$阶、2 阶,以致更高阶的差分近似模型.但是,对于通常的应用,一阶近似模型已经完全足够了.

下面的内容,有助于我们对得到更精确的随机差分模型近似的认识,其核心思想就是多次运用 Ito 公式.

8.3　Ito 方法的解 ξ_t 的光滑函数 $f(\xi_t)$在时刻 t 附近的随机 Taylor 展开

对于 Ito 过程 ξ_t,我们考虑 $f(\xi_t)$在 t 附近的近似计算.

为了求得较好的近似,我们要在 $t_m<t\leqslant t_m+\Delta t_m$ 时用 Ito 过程的函数 $f(\xi_t)$作**随机 Taylor 展开式**,它是 8.2 节的思路的推广,而 8.2 节则可以看成 $f(x)=x$ 的情形.

由 Ito 公式

$$f(\xi_t)=f(\xi_{t_m})+\int_{t_m}^{t}\left[b(\xi_s)f'(\xi_s)+\frac{1}{2}\sigma^2(\xi_s)f''(\xi_s)\right]\mathrm{d}s+\int_{t_m}^{t}\sigma(\xi_s)f'(\xi_s)\mathrm{d}B_s,$$

它的$\frac{1}{2}$阶**随机 Taylor 展开式**则是在上式中,将被积的随机过程 ξ_s 简单地用(它在时刻 t_m 处的随机变量)ξ_{t_m} 替代后,再把两者的差项都归结为一个余项,这就是

$$f(\xi_t)=f(\xi_{t_m})+\left[b(\xi_{t_m})f'(\xi_{t_m})+\frac{1}{2}\sigma^2(\xi_{t_m})f''(\xi_{t_m})\right]\int_{t_m}^{t}\mathrm{d}s+\sigma(\xi_{t_m})f'(\xi_{t_m})\int_{t_m}^{t}\mathrm{d}B_s+R,$$

其中 R 是一个随机的余项.

较高精确度的是一阶随机 Taylor 展开,为了得到它,把上式改写为

$$f(\xi_t)=f(\xi_{t_m})+\left[b(\xi_{t_m})f'(\omega_{t_m})+\frac{1}{2}\sigma^2(\xi_{t_m})f''(\xi_{t_m})\right]\int_{t_m}^{t}\mathrm{d}s+\sigma(\xi_{t_m})f'(\xi_{t_m})\int_{t_m}^{t}\mathrm{d}B_s$$

$$+\int_{t_m}^{t}[\sigma(\xi_s)f'(\xi_s)-\sigma(\xi_{t_m})f'(\xi_{t_m})]\mathrm{d}B_s+R_1.$$

对于乘积复合函数$(\sigma\cdot f')(\xi_s)$在应用 Ito 公式,我们发现$\sigma(\xi_s)f'(\xi_s)-\sigma(\xi_{t_m})f'(\xi_{t_m})$中含 Ito 积分的项只有$\int_{t_m}^{s}\sigma(\xi_u)(\sigma\cdot f')'(\xi_u)\mathrm{d}B_u$,它的$\frac{1}{2}$阶近似为

$$\sigma(\xi_{t_m})(\sigma\cdot f')'(\xi_{t_m})(B_s-B_{t_m}),$$

其他部分都高于$\frac{1}{2}$阶.于是在只考虑一阶近似时,可以将展开式写为

$$f(\xi_t)=f(\xi_{t_m})+\left[b(\xi_{t_m})f'(\xi_{t_m})+\frac{1}{2}\sigma^2(\xi_{t_m})f''(\xi_{t_m})\right]\int_{t_m}^{t}\mathrm{d}s+\sigma(\xi_{t_m})f'(\xi_{t_m})\int_{t_m}^{t}\mathrm{d}B_s$$
$$+\sigma(\xi_{t_m})(\sigma\cdot f')'(\xi_{t_m})\int_{t_m}^{t}\int_{t_m}^{s}\mathrm{d}B_u\mathrm{d}B_s+R_1. \tag{8.10}$$

这就回到一阶随机 Taylor 展开式.

因此,对于随机微分方程中的被积随机过程用本节的方法展开到足够的阶,就可以得到更为精确的随机差分近似模型.

8.4 差分近似模型的另一种改进途径——一阶随机 Runge-Kutta 模型

用一阶随机 Taylor 展开的思想作差分近似的缺点是,需要计算扩散系数$\sigma(x)$的微商.在实践中$\sigma(x)$可能没有数学表达式.避免使用其微商的一个思路是:借助于常微分方程近似计算的思想,用**一阶随机 Runge-Kutta 模型**.也就是用随机差分模型

$$\xi_{t_{m+1}}^{(n)}=\xi_{t_m}^{(n)}+b(\xi_{t_m}^{(n)})\Delta t_m+\sigma(\xi_{t_m}^{(n)})\xi_m+\frac{1}{2}\left\{\frac{1}{\sqrt{\delta}}[\sigma(\xi_{t_m}^{(n)}+\sigma(\xi_{t_m}^{(n)})\sqrt{\delta})-\sigma(\xi_{t_m}^{(n)})]\right\}(\xi_m^2-\Delta t_m),$$
$$\{\zeta_m\}\text{ 独立同分布},\quad \xi_m\sim N(0,\delta),\quad \xi_{t_0}^{(n)}=x, \tag{8.11}$$

在$n\to+\infty$时,理论上可以证明,以上几个模型得到的"近似解"确实是在平均意义下收敛到随机微分方程的唯一的解.

以上讨论的是取值于实直线$\mathbb{R}^1$的随机微分方程.对于多变量情形,即取值于$\mathbb{R}^d$的随机微分方程,可以完全类似地得到随机 Taylor 近似及随机 Runge-Kutta 模型.本书不再赘述.

第 9 章　随机微分方程在金融模型中的应用

9.1　金融术语与基本假定

在金融理论中有一个重要的概念就是套利，即人们受利益的驱使，总是在寻找套利的机会. 套利的直观含义是：在开始时并无资金，经过资本的市场运作后(包括利用借贷)，变成为非负的(随机)资金，而且以正概率具有正的资金，也就是通常所说的"空手套白狼"的机会.

金融市场的基本假定——无套利假定

研究金融市场常常设置一个基本假定，称为**无套利原则**，也称**套利原则**，就是假定正常运行的市场没有套利机会. 事实上，因为在出现套利机会时，大量的投机者就会涌向市场进行套利，于是经过一个相对短的时期的"混乱"后，市场就会重返"正常"，即回复到无套利状态. 在金融衍生证券的定价理论中并不讨论这段短混乱时期，因此，在研究中，普遍地设置无套利假定.

定义 9.1(可行市场)　满足无套利假定的市场称为**可行市场**.

定义 9.2(套期)　直观地说，以持有某些有价证券的组合来抵消某种金融衍生证券所带来的风险，称为**套期**. 这种套期事实上是**完全套期**. 如果只抵消了部分风险，则称为**部分套期**.

一种金融证券每单位在 t 时刻的价格为 S_t，通常地 $\{S_t, t\geqslant 0\}$ 是一个随机过程，常见的情形是一个随机微分方程的解，或者是一个离散时间的随机序列 $\{S_n, n\geqslant 0\}$. 一般地，将

$$\frac{1}{S_t}\lim_{h\to 0}\frac{E([S_{t+h}-S_t]\mid S_u, u\leqslant t)}{h}$$

和

$$\frac{1}{S_t}\lim_{h\to 0}\frac{\sqrt{\mathrm{var}([S_{t+h}-S_t]\mid S_u, u\leqslant t)}}{h}$$

分别称为这个金融证券的**收益率**(yield)和**波动率**(volatility). 前者显示单位时间单位价格的平均收益，而后者表示单位时间单位价格的平均波幅变动，它们是描述证券的重要指标. 在一般情形，它们都是随机的且随时间变化的量，因此也可能是随机过程.

定义 9.3(欧式期权，欧式未定权益)　以一种证券为对象(称为**标的变量**，underlying

variable)的**欧式看涨期权**(european call option),是指在 $t=0$ 时甲方(一般为证券公司)与乙方的一个合约,按此合约规定乙方有一个权利,能在时刻 T 以价格 K(它称为**执行价格**,striking price,executable price)从甲方买进一批(一般为 100 份)这种证券,而且如果时间 T 时的市场价格 S_T 低于 K,乙方可以不买.按照这个和约只要在时间 T 时证券的市场价格 S_T 高于 K,乙方就得利.于是,乙方在时刻 T 可净得随机收益 $X_T=(S_T-K)^+$,其中 $a^+=\begin{cases}a, & a>0,\\0, & a\leqslant 0\end{cases}$ 是实数 a 的正部.这种合约称为**期权**(option).这种只允许乙方在最终时刻 T 作出选择的期权,称为**欧式期权**.这时乙方希望 S_T 尽量大,以便更多地获利.也就是有选择权的乙方盼望股票上涨,所以称为**看涨期权**,或者**买权**(call option).同样,这个合约能给乙方带来 X_T 的随机收益,这也需要乙方在 $t=0$ 时刻用钱从甲方购买.这个合约在 $t=0$ 时刻的价格,也称为它的**贴水**或**保证金**(premium).研究金融衍生证券的第一个问题,是如何确定这个合约在时刻 $t<T$ 的价格.

另一种相反的情况是,如果 $t=0$ 时甲方(一般为证券公司)卖给乙方如下的合约,此合约规定乙方有一个权利,使乙方在时刻 T 以价格为 K 卖给甲方一批(一般也为 100 份)这种证券,如果在时刻 T 时的市场价格 S_T 高于 K,乙方可以不卖.按照这个和约只要在时间 T 时证券的市场价格 S_T 低于 K,卖方就得利.于是,乙方在时刻 T 能净得随机收益 $X_T=(K-S_T)^+$.这也是一种只允许乙方在最终时刻 T 作出选择的期权,所以也是一种欧式期权,此时乙方盼望 S_T 尽量小,以便有更多的获利.也就是,乙方盼望股票下跌,所以称为**看跌期权**,或者**卖权**(put option).同样由于这个合约也能给乙方带来 X_T 的收益,也就需要乙方在 $t=0$ 时刻用钱从甲方购买.这个合约在 $t=0$ 时刻的价格,也称为它的贴水或保证金.

一些大公司常把以本公司的股票为标的证券的期权,作为其雇员收入的一部分,以将雇员与公司利益更紧密地联系起来.

比看涨期权与看跌期权更为一般的欧式期权是:甲方卖给乙方一个由证券组合组成的一个合约,此合约能在 T 时刻给乙方带来随机收益 $f(S_T)$(称为**欧式未定权益**,European contingent claim),我们也要给出这个合约在时刻 $t<T$ 的价格.看涨期权和看跌期权都是一种特殊的未定权益,分别对应于 $f(x)=(x-K)^+$ 和 $f(x)=(K-x)^+$.

另一方面,如果允许乙方在最终时刻 T 前的任意时刻都可以作出选择的未定权益则称为**美式未定权益**(American contingent claim),一般地,美式未定权益依赖于一个双变量函数 $f(t,x)$.在时刻 t 执行此权益的收益是 $f(t,S_t)$.即使在不显含 t 的情形,即 $f(t,S_t)=f(S_t)$ 的情形美式未定权益也应该比欧式未定权益的价格高,因为它给予乙方更多的选择自由.

基本假定 为了讨论简单,通常假定市场是可行的,而且在下述意义下是**无摩擦的**,即无税收、无交易费,允许卖空(借贷证券与现金),银行的存贷利率是一样的.再假定此未定权益的持有人是小投资者,并且是**自融资的**,即在整个过程中,持有人既没有添入资金,也没有抽走资金.

9.2 Black-Scholes 模型及其欧式未定权益的定价

9.2.1 Black-Scholes 模型

假设 S_t 满足以下的 Black-Scholes 模型：

$$dS_t = S_t(\mu dt + \sigma dB_t). \tag{9.1}$$

直观地可以看出，其中常数 $\mu,\sigma(>0)$分别恰是证券的**收益率**与**波动率**.

假定当前的银行利率是非随机的常数 r（或较为一般些，是时间的决定性函数$r(t)$）. 存银行的钱是无风险的，由银行利率为 r（或 $r(t)$，这在数学处理上并未增加任何困难），时刻 $t=0$ 的存款 S_0^0 元到时刻 t 的价值应为

$$S_t^0 = S_0^0 e^{rt}, \quad 即 \quad dS_t^0 = rS_t^0 dt. \tag{9.2}$$

或

$$S_t^0 = S_0^0 e^{\int_0^t r(s)ds}, \quad 即 \quad dS_t^0 = r(t)S_t^0 dt.$$

无套利下的欧式看涨与看跌期权的平权关系

定义 9.4（远期合约） 未定权益为 S_T 的欧式权益，称为在时刻 T 成熟的**远期合约**. 显见远期合约在时刻 $t(<T)$的价格就应该为证券的即时价格 S_t，因此，它也是一种欧式未定权益.

命题 9.1（平权关系） 如果将看涨期权、看跌期权、远期合约在时刻 $t(<T)$的价格分别记为 C_t, P_t, F_t. 那么，在无套利假定下显然有

$$C_t - P_t = F_t - Ke^{-r(T-t)}. \tag{9.3}$$

这个关系式称为**欧式看涨-看跌期权的平权关系**（call-put parity).

证明 只需注意$(x-K)^+ - (K-x)^+ = x-K$，即未定权益有等式：

$$(S_T - K)^+ - (K - S_T)^+ = S_T - K.$$

平权关系说明买进一张（在金融中称为**多头**一张）在 T 到期的执行价格为 K 的看涨期权与卖出一张（在金融中称为**空头**一张）相应的看跌期权，就相当于买进一张远期合约与卖出一张在时刻 T 到期的额度为 K 的银行存款.

有了平权关系，在欧式看涨期权与看跌期权中只要知道一个的价格，就立刻可以得到另一个的价格.

9.2.2 Black-Scholes 模型的欧式未定权益的定价的套期方法，Black-Scholes 微分方程的推导与求解

(1) 欧式未定权益的定价的 Black-Scholes 偏微分方程及其推导

满足 Black-Scholes 模型(9.1)的风险标的资产（即证券）S_t 是随机微分方程的解，因

此它是 Markov 过程.于是,欧式未定权益 $f(S_T)$在时刻 $t(<T)$的价格(当然是在平均的意义下)只依赖标的证券在时刻 t 当前的价格 S_t,因此我们可以将此价格记为 $V(t,S_t)$.下面我们按 Black-Scholes 在 1973 年的经典论文中的套期思想,来推导价格函数 $V_t \stackrel{\text{def}}{=\!=} V(t,x)$ 满足的微分方程.首先,利用 Ito 公式有

$$\mathrm{d}V(t,S_t) = \left[\frac{\partial V}{\partial x}\sigma S_t \mathrm{d}B_t + \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}\right)\mathrm{d}t\right]_{x=S_t}.$$

套期思想是,假定在时刻 t 甲方为了保护卖出未定权益的风险,虚拟地待定购进 Δ 份标的证券.那么,将卖出的未定权益合约得到的价格 V_t 与购进的标的证券的花费 ΔS_t 合起来考虑,甲方的总资产是

$$N_t \stackrel{\text{def}}{=\!=} V_t - \Delta S_t.$$

由自融资原则,到时刻 $t+\mathrm{d}t$,这个资产就变为 $N_{t+\mathrm{d}t} \stackrel{\text{def}}{=\!=} V_{t+\mathrm{d}t} - \Delta S_{t+\mathrm{d}t}$.即

$$\mathrm{d}N_t = \mathrm{d}V - \Delta \mathrm{d}S_t = \left[\left(\frac{\partial V}{\partial x} - \Delta\right)\sigma S_t \mathrm{d}B_t + \left(\frac{\partial V}{\partial t} + \mu S_t\left(\frac{\partial V}{\partial x} - \Delta\right) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}\right)\mathrm{d}t\right]_{x=S_t}.$$

这样,只要取

$$\Delta = \frac{\partial V}{\partial x}(t,S_t), \tag{9.4}$$

就得到

$$\mathrm{d}N_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}\right)_{x=S_t} \mathrm{d}t. \tag{9.5}$$

再由无套利假定可知 N_t 必须是无风险资产,即

$$\mathrm{d}N_t = rN_t \mathrm{d}t. \tag{9.6}$$

事实上,如果 $\mathrm{d}N_t > rN_t\mathrm{d}t$,那么,甲可以在时刻 t 从银行借贷并投资于上述组合 N_t,而在 $t+\mathrm{d}t$ 时刻得利 $\mathrm{d}N_t$ 后立刻偿还银行 $rN_t\mathrm{d}t$,从而净得 $\mathrm{d}N_t - rN_t\mathrm{d}t$.在另一种情形,如果 $\mathrm{d}N_t < rN_t\mathrm{d}t$,那么,甲可以在时刻 t 卖空上述组合 N_t,将钱存入银行,到 $t+\mathrm{d}t$ 时刻得利 $rN_t\mathrm{d}t$,并且再购回卖空的组合,由此也净得利 $rN_t\mathrm{d}t - \mathrm{d}N_t$.所以,这两种情形都发生了套利,与无套利假定矛盾.可见只能有 $\mathrm{d}N_t = rN_t\mathrm{d}t$.

将(9.6)式中的 N_t 的表达式代入(9.5)式,就得到 $V(t,x)$应满足的如下的 Black-Scholes 偏微分方程:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2} + rx\frac{\partial V}{\partial x} - rV = 0, \tag{9.7}$$

并且附加如下的终端条件

$$V(T,x) = f(x) \quad (\text{因为 } V(T,S_T) = f(S_T)). \tag{9.8}$$

我们只需求得此方程的解 $V(t,x)$,就得到了未定权益 $f(S_T)$在时刻 $t<T$ 的价格 $V(t,S_t)$.而在时刻 $t=0$ 贴水就是 $V(0,S_0)$.

(2) 求解 Black-Scholes 微分方程的方法

先令 $t'=T-t$ 将终值问题化成初值问题,再令 $x'=\log x,\widetilde{V}(t',x')=V(t,x)$,并利用

$$\widetilde{V}_{x'}=xV_x,\quad \widetilde{V}_{x'x'}=(xV_x)_{x'}=x(xV)_x=x^2V_{xx}+xV_x,\quad \widetilde{V}_{t'}=-V_t,$$

将方程(9.7)化简为

$$-\widetilde{V}_{t'}+\frac{1}{2}\sigma^2\widetilde{V}_{x'x'}+\left(r-\frac{\sigma^2}{2}\right)\widetilde{V}_{x'}-r\widetilde{V}=0,\tag{9.9}$$

$$\widetilde{V}(0,x')=(V(T,x)=f(x)=)f(e^{x'}).\tag{9.10}$$

这个方程的系数不依赖 x',它是常系数偏微分方程. 然后,我们作变换

$$\widetilde{V}(t',x')=e^{\alpha x'+\beta t'}U(t',x').\tag{9.11}$$

只要在此变换中选取合适的常数 α,β,就可以消灭方程中含 U 和 $U_{x'}$ 的项,具体步骤如下.

第一步: 要求 $U_{x'}$ 的系数为 0,由此得到

$$\alpha=-\frac{r-\frac{\sigma^2}{2}}{\sigma^2}=\frac{1}{2}-\frac{r}{\sigma^2},\tag{9.12}$$

第二步: 再要求 U 的系数为 0,得到

$$\beta=-\left(\frac{1}{2}\sigma^2\alpha^2+r\right).\tag{9.13}$$

这时(9.9)式,(9.10)式就简化为传热方程的初值问题

$$U_{t'}=\frac{\sigma^2}{2}U_{x'x'},\tag{9.14}$$

$$U(0,x')=e^{-\alpha x'}f(e^{x'}).\tag{9.15}$$

由此可以用初等偏微分方程中的经典方法,即用 **Gauss 核**(Brown 运动的转移密度函数也称为 Gauss 核)的积分给出其解为

$$\begin{aligned}U(t',x')&=\frac{1}{\sqrt{2\pi t'}\sigma}\int U(0,y)e^{-\frac{(y-x')^2}{2t'\sigma^2}}dy\\&=\frac{1}{\sqrt{2\pi t'}\sigma}\int e^{-\alpha y}f(e^y)e^{-\frac{(y-x')^2}{2t'\sigma^2}}dy.\end{aligned}$$

再作变换 $y-x'=y'$,便得

$$U(t',x')=\frac{1}{\sqrt{2\pi t'}\sigma}\int f(e^{y'+x'})e^{-\frac{y'^2}{2t'\sigma^2}}e^{-\alpha(x'+y')}dy'=\frac{e^{-\alpha x'}}{\sqrt{2\pi t'}\sigma}\int f(xe^{y'})e^{-\frac{y'^2+2t'\sigma^2\alpha y'}{2t'\sigma^2}}dy'.$$

注意积分号内指数项的分子为 $(y'+t'\sigma^2\alpha)^2-t'^2\sigma^4\alpha^2$,作变换 $\dfrac{y'+t'\sigma^2\alpha}{\sigma\sqrt{t'}}=z$,就得

$$U(t',x')=\frac{e^{-\alpha x'}}{\sqrt{2\pi}}\int f(xe^{\sigma\sqrt{t'}z-t'\sigma^2\alpha})e^{-\frac{z^2}{2}}e^{\frac{t'\sigma^2\alpha^2}{2}}dz$$

$$= \frac{e^{-\alpha x'} e^{-(\beta t'+rt')}}{\sqrt{2\pi}} \int f(x e^{\sigma z \sqrt{T-t}-(T-t)\sigma^2 \alpha}) e^{-\frac{z^2}{2}} dz.$$

最后用 $V(t,x)=\widetilde{V}(t',x')=e^{\alpha x'+\beta t'}U(t',x')$，得到 $V(t,x)$ 的如下面定理所述的明显表达式.

定理 9.2

$$V(t,x) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int f(x e^{\sigma z \sqrt{T-t}+(T-t)\left(r-\frac{\sigma^2}{2}\right)}) e^{-\frac{z^2}{2}} dz. \tag{9.16}$$

于是欧式未定权益 $f(S_T)$ 在时刻 $t(<T)$ 的价格为 $V(t,S_t)$，而在合约开始时刻的贴水为

$$V(0,S_0) = \frac{e^{-rT}}{\sqrt{2\pi}} \int f(S_0 e^{\sigma z \sqrt{T}+T\left(r-\frac{\sigma^2}{2}\right)}) e^{-\frac{z^2}{2}} dz. \tag{9.17}$$

而用以套期标的证券的数量，则由(9.4)式给出.

定义 9.5(Black-Scholes 中性模型)　从(9.16)式可以看出，欧式未定权益的定价不依赖风险证券的收益率 μ. 而代替它的则是银行利率 r. 这就启示我们，在利用 Black-Scholes 模型求未定权益的定价时，风险证券的价格模型可以改用

$$dS_t = S_t(r dt + \sigma dB_t). \tag{9.18}$$

这样的模型称为 **Black-Scholes 风险中性模型**.

由于风险中性模型(9.18)与收益率为 μ 时的模型不一样，我们将对应于这个风险中性模型所取的概率记为 P^*（相应的期望记为 E^*），称为**风险中性概率**.

$\frac{\mu-r}{\sigma}$ 是收益率超出银行利率的部分关于波动率的倍数，在金融中称为**风险的市场价格**.

9.2.3　风险中性概率方法

设在风险中性模型下，风险证券在 t 时刻的价格记为 S_t. 它关于银行利率 r 的折现价记为 $\widetilde{S}_t \stackrel{\text{def}}{=} e^{-rt}S_t$（在初始时刻它的等价价格，就是将一切资产都折合到初始时刻考虑，以便比较）. 对它用 Ito 公式得

$$d\widetilde{S}_t = d(e^{-rt}S_t) = -re^{-rt}S_t dt + e^{-rt}S_t(r dt + \sigma dB_t) = \sigma \widetilde{S}_t dB_t. \tag{9.19}$$

由此可见对风险中性概率 P^* 而言，风险证券的折现价 $\widetilde{S}_t$ 是随机微分方程(9.19)的解，由例 6.5，这个方程有唯一解，这个解是

$$\widetilde{S}_t = \widetilde{S}_0 e^{-\frac{\sigma^2}{2}t+\sigma B_t}. \tag{9.20}$$

这正是 Brown 运动 B_t 的一个指数鞅. 又因为鞅在时间进行中体现了公平的（平均地）发展，时刻 T 的未定权益 $f(S_T)$ 在初始时刻的折现价是 $\tilde{f} \stackrel{\text{def}}{=} e^{-rT}f(e^{rT}\widetilde{S}_T)$，而鞅性质应该体

现为：时刻 t 的价格在初始时刻的折现价 $\widetilde{V}_t \stackrel{\text{def}}{=} e^{-rt}V_t$ 应该是 $E^*(\tilde{f}|B_s,s\leqslant t)$，即

$$\widetilde{V}_t = E^*(\tilde{f} \mid B_s, s \leqslant t). \tag{9.21}$$

于是

$$\begin{aligned} V_t &= e^{rt}E^*(e^{-rT}f(e^{rT}\widetilde{S}_T) \mid B_s, s \leqslant t) \\ &= e^{-r(T-t)}E^*[f(S_t e^{\sigma(B_T-B_t)+\left(r-\frac{\sigma^2}{2}\right)(T-t)}) \mid B_s, s \leqslant t]. \end{aligned} \tag{9.22}$$

再利用条件期望的公式，并令 $\eta \stackrel{\text{def}}{=} \dfrac{B_T-B_t}{\sqrt{T-t}}$，则对应于概率 P^* 有 $\eta \sim N(0,1)$. 故而

$$\begin{aligned} V_t &= [e^{-r(T-t)}E^* f(x e^{\sigma(B_T-B_t)})]_{x=S_t \exp\left[\left(r-\frac{\sigma^2}{2}\right)(T-t)\right]} \\ &= [e^{-r(T-t)}E^* f(x e^{\sigma\sqrt{T-t}\eta})]_{x=S_t \exp\left[\left(r-\frac{\sigma^2}{2}\right)(T-t)\right]} \\ &= e^{-r(T-t)}\frac{1}{\sqrt{2\pi}}\int f(S_t e^{\left(r-\frac{\sigma^2}{2}\right)(T-t)+\sigma\sqrt{T-t}u})e^{-\frac{u^2}{2}}du \stackrel{\text{def}}{=} V(t,S_t). \end{aligned}$$

$V(t,x)$ 正好是带终端条件的 Black-Scholes 偏微分方程的解(9.16).

推论 9.3 对于远期合约 $f(x)=x$，由定理 9.2 得到它在时刻 $t(t<T)$ 的价格为 $F(t,S_t)=S_t$，恰好就是标的风险证券的市价.

推论 9.4 对于欧式看涨期权(对应于 $f(x)=(x-K)^+$)，它在时刻 $t(t<T)$ 的价格 $F(t,S_t)$ 的公式可以简化为以下的 Black-Scholes-Merton 公式：

$$F(t,x) = x\Phi(d_1(x)) - Ke^{-r(T-t)}\Phi(d_2(x)),$$

其中

$$d_1(x) = \frac{\log\left(\frac{x}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2(x) = \frac{\log\left(\frac{x}{K}\right)+\left(r-\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}},$$

$\Phi(x) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x} e^{-\frac{u^2}{2}}du$（在通常的金融文献中记为 $N(x)$）是为标准正态分布的分布函数. 此时应该套期的风险证券的数量为

$$\Delta_t = \left.\frac{\partial F}{\partial x}\right|_{x=S_t} = [\Phi(d_1(x)) + x e^{-\frac{d_1(x)^2}{2}}d_1'(x) - Ke^{-r(T-t)}e^{-\frac{d_2(x)^2}{2}}d_2'(x)]_{x=S_t} = \Phi(d_1(S_t)).$$

推论 9.5 由平权关系便得到欧式看跌期权(对应于 $f(x)=(K-x)^+$)在时刻 $t(t<T)$ 的价格

$$F(t,x) = Ke^{-r(T-t)}\Phi(-d_2(x)) - x\Phi(-d_1(x)).$$

此时应该套期的风险证券的数量为

$$\Delta_t = \left.\frac{\partial F}{\partial x}\right|_{x=S_t} = \cdots = -\Phi(-d_1(S_t)).$$

注 如果不用风险中性的 Black-Scholes 模型,而是用带收益率 μ 的 Black-Scholes 模型,那么类似地用 Ito 公式,可得证券的折现价格$\widetilde{S}_t$ 满足的方程为

$$\mathrm{d}\widetilde{S}_t = \widetilde{S}_t\sigma\left(\mathrm{d}B_t + \frac{\mu - r}{\sigma}\mathrm{d}t\right).$$

作 Girsanov 变换 $B^* = B_t + \frac{\mu - r}{\sigma}t$,则对于$\mathcal{F}$中的任意事件 A,只要它的信息完全可由$(B_s: s \leqslant t)$确定,就定义它的一个新概率 P^* 为

$$P^*(A) = E(I_A \mathrm{e}^{-\frac{\mu-r}{\sigma}B_t - \frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2 t}).$$

于是由 Girsanov 定理(定理 7.16),概率 P^* 下,B^* 是一个 Brown 运动. 从而重新得到 $\mathrm{d}\widetilde{S}_t = \widetilde{S}_t\sigma\,\mathrm{d}B_t^*$. 这正好是风险中性的 Black-Scholes 模型. 而且概率 P^* 就是风险中性概率. 这样用另一个方法得到同样的结果.

*9.2.4 币值单位与随机折现因子方法

定义 9.6 设某个风险证券在时刻 t 的价格为 S_t. 如果存在另一个正值随机过程 M_t 使得$\widetilde{S}_t \stackrel{\text{def}}{=} \frac{S_t}{M_t}$是一个鞅,则称 M_t 为证券 S_t 的**币值单位**(更一般地,在理论上只要$\widetilde{S}_t$ 是一个局部鞅,也就是存在一个$(\widetilde{S}_t)$停时序列 $\tau_n \uparrow +\infty$,使对于固定的 n,$\{\widetilde{S}_{t\wedge\tau_n}: t \geqslant 0\}$是鞅. 这里, 鞅和局部鞅都代表在统计平均意义下不随时间增值的资金流).

从直观看,资本的价值是随着时间增值的,币值单位 M_t 就体现了该证券的价格的时间增值,即在 $t=0$ 时的 1 元钱,经过单个证券 S_t 的市场因素的作用,在时刻 t 的实际价值相当于 M_t 元. 又因为市场是随机的,所以币值单位也是随机的量(随机过程),这就体现了一元钱的时间价值.

定义 9.7 币值单位 M_t 的倒数

$$N_t \stackrel{\text{def}}{=} \frac{1}{M_t} \tag{9.23}$$

称为**随机折现因子**,因为 N_t 是 1 除以币值单位,所以它就相当于把随机因素考虑进去以后的"随机折现因子".

Black-Scholes 模型的币值单位是什么呢?对此我们有下面的命题.

命题 9.6 假定证券在时刻 t 的价格 S_t 满足 Black-Scholes 模型

$$\mathrm{d}S_t = S_t(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}B_t). \tag{9.24}$$

那么

$$M_t = \mathrm{e}^{rt}L_t, \tag{9.25}$$

就是 S_t 的币值单位,其中 r 是银行利率,其中

$$L_t \stackrel{\text{def}}{=} \mathrm{e}^{\frac{\mu-r}{\sigma}B_t + \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}t} \stackrel{\text{def}}{=} \mathrm{e}^{G(t,B_t)}. \tag{9.26}$$

证明 对于 L_t 用 Ito 公式得到

$$\begin{aligned}\mathrm{d}L_t &= \mathrm{e}^{G(t,B_t)}\mathrm{d}G(t,B_t) + \frac{1}{2}\mathrm{e}^{G(t,B_t)}(\mathrm{d}G(t,B_t))^2 \\ &= \mathrm{e}^{G(t,B_t)}\left[\frac{\mu-r}{\sigma}\mathrm{d}B_t + \left(\frac{\mu-r}{\sigma}\right)^2\mathrm{d}t\right].\end{aligned}$$

于是 M_t 及 $N_t=\dfrac{1}{M_t}$满足的随机微分分别为

$$\begin{aligned}\mathrm{d}M_t &= r\mathrm{e}^{rt}\mathrm{d}tL_t + \mathrm{e}^{rt}\mathrm{d}L_t = rM_t\mathrm{d}t + M_t\left[\frac{\mu-r}{\sigma}\mathrm{d}B_t + \left(\frac{\mu-r}{\sigma}\right)^2\mathrm{d}t\right] \\ &= M_t\left(\left[r+\left(\frac{\mu-r}{\sigma}\right)^2\right]\mathrm{d}t + \frac{\mu-r}{\sigma}\mathrm{d}B_t\right). \\ \mathrm{d}N_t &= -\frac{1}{M_t^2}\mathrm{d}M_t + \frac{1}{M_t^3}(\mathrm{d}M_t)^2 \\ &= -\frac{1}{M_t}\left(\left[r+\left(\frac{\mu-r}{\sigma}\right)^2\right]\mathrm{d}t + \frac{\mu-r}{\sigma}\mathrm{d}B_t\right) + \frac{1}{M_t^3}M_t^2\left(\frac{\mu-r}{\sigma}\right)^2\mathrm{d}t \\ &= -N_t\left(r\mathrm{d}t + \frac{\mu-r}{\sigma}\mathrm{d}B_t\right).\end{aligned}$$

即 N_t 所满足的随机微分方程为

$$\mathrm{d}N_t = N_t\left(-r\mathrm{d}t - \frac{\mu-r}{\sigma}\mathrm{d}B_t\right). \tag{9.27}$$

再用 Ito 公式得到

$$\begin{aligned}\mathrm{d}(S_tN_t) &= S_t\mathrm{d}N_t + N_t\mathrm{d}S_t + \mathrm{d}S_t\mathrm{d}N_t \\ &= S_tN_t\left(-r\mathrm{d}t - \frac{\mu-r}{\sigma}\mathrm{d}B_t\right) + N_tS_t(\mu\mathrm{d}t + \sigma\mathrm{d}B_t) + S_t\sigma N_t\left(-\frac{\mu-r}{\sigma}\right)\mathrm{d}t \\ &= S_tN_t\left(\sigma - \frac{\mu-r}{\sigma}\right)\mathrm{d}B_t.\end{aligned}$$

由此可知,$\dfrac{S_t}{M_t}=S_tN_t$ 是一个指数鞅 (严格地,是局部鞅). 命题得证.

既然$\widetilde{S}_t \overset{\mathrm{def}}{=} \dfrac{S_t}{M_t}$ 随时间的发展体现了一个“公平博弈”,由此可得到,一般的欧式未定权益 X_T 与它们在 $t(<T)$的定价 V_t 在时刻 $t=0$ 的折现$\widetilde{V}_t \overset{\mathrm{def}}{=} \dfrac{V_t}{M_t}$,$\widetilde{X}_T \overset{\mathrm{def}}{=} \dfrac{X_T}{M_T}$之间应该有

$$\widetilde{V}_t = E(\widetilde{X}_T \mid B_s;\ s\leqslant t).$$

从而得到更为一般的定价公式

$$V_t = M_tE\left(\frac{X_T}{M_T}\middle|B_s, s\leqslant t\right) = \mathrm{e}^{rt}L_tE\left(X_T\mathrm{e}^{-rT}\frac{1}{L_T}\middle|B_s, s\leqslant t\right). \tag{9.28}$$

9.2.5 时变的 Black-Scholes 模型

设风险证券在 t 时刻的价格 S_t 满足

$$dS_t = S_t(\mu(t)dt + \sigma(t)dB_t), \tag{9.29}$$

其中 $\mu(t),\sigma(t)(>0)$是决定性的函数,分别代表证券的时变收益率与波动率.基于这种模型的证券衍生的欧式未定权益的定价,用上述三种方法中的任意一种都可以最后得到

$$V(t,x) = e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} \int f\left(x e^{\int_t^T \left(r(s)-\frac{\sigma(s)^2}{2}\right)ds + \frac{z}{\sqrt{T-t}}\int_t^T \sigma(s)ds}\right) e^{-\frac{z^2}{2}} dz. \tag{9.30}$$

从而,贴水为

$$V(0,S_0) = e^{-rT} \frac{1}{\sqrt{2\pi}} \int f(S_0 e^{\int_0^T \left(r(s)-\frac{\sigma(s)^2}{2}\right)ds + \frac{z}{\sqrt{T}}\int_0^T \sigma(s)ds}) e^{-\frac{z^2}{2}} dz.$$

9.3 二叉模型与 Black-Scholes 模型的二叉近似

9.3.1 二叉模型(中性概率存在的条件,欧式权益的定价)

(1) 二叉模型与中性概率存在的条件

二叉模型又称 **Cox-Ross-Rubinstein 模型**,它考虑时间离散情形,即假定风险证券价格的增长为

$$S_{n+1} = (1+\eta_n)S_n, \tag{9.31}$$

其中$\{\eta_n\}$是独立同分布的随机变量列,且

$$1+\eta_n \sim \begin{pmatrix} a & b \\ p & 1-p \end{pmatrix} \quad (0 < a < b, 0 < p < 1),$$

即

$$\frac{S_{n+1}}{S_n} \sim \begin{pmatrix} a & b \\ p & 1-p \end{pmatrix}. \tag{9.31$'$}$$

假设银行利率为常数 R(注意离散时间的利率与连续实际的利率之间有一个折合关系),即存入银行的资金的增长为

$$S_{n+1}^0 = (1+R)S_n^0. \tag{9.32}$$

于是,用二叉模型建模的证券在时刻 $t=0$ 的折现价格为

$$\widetilde{S}_n \overset{\text{def}}{=\!=} \frac{1}{(1+R)^n} S_n. \tag{9.33}$$

二叉模型对应于离散的 Black-Scholes 模型.但是它比 Black-Scholes 模型更适合于实际拟合或模拟计算.

受到 Black-Scholes 模型的启示，无套利相当于对于$\frac{S_{n+1}}{S_n}$存在一个如下的概率分布

$$\begin{pmatrix} a & b \\ p^* & 1-p^* \end{pmatrix}, \tag{9.34}$$

（称为风险中性的概率分布）使风险证券的折现价格$\widetilde{S}_n$关于这个分布是鞅列，即满足

$$E^*(\widetilde{S}_{n+1} \mid \widetilde{S}_n, \cdots, \widetilde{S}_0) = \widetilde{S}_n,$$

其中 P^* 表示$\frac{S_{n+1}}{S_n}$的分布是(9.34)对应于事件类$\mathcal{F}$上定义的风险中性概率 $P^*(\cdot)$，而概率 p^* 则称为公平概率.

定理 9.7　当且仅当 $a<1+R<b$ 时，公平概率 p^* 存在唯一，在条件成立时有

$$p^* = \frac{b-(1+R)}{b-a}. \tag{9.35}$$

证明　由于$\left\{\frac{S_{n+1}}{S_n}\right\}$的独立同分布性推出$\frac{\widetilde{S}_{n+1}}{\widetilde{S}_n}=\frac{1}{1+R}\frac{S_{n+1}}{S_n}$与$\{\widetilde{S}_n,\cdots,\widetilde{S}_0\}$是相互独立的. 因此

$$\widetilde{S}_n = E^*(\widetilde{S}_{n+1} \mid \widetilde{S}_n, \cdots, \widetilde{S}_0) = E^*\left(\frac{\widetilde{S}_{n+1}}{\widetilde{S}_n}\widetilde{S}_n \mid \widetilde{S}_n, \cdots, \widetilde{S}_0\right) = \widetilde{S}_n E^*\left(\frac{\widetilde{S}_{n+1}}{\widetilde{S}_n}\right).$$

从而有

$$1 = E^*\left(\frac{\widetilde{S}_{n+1}}{\widetilde{S}_n}\right) = \frac{a}{1+R}p^* + \frac{b}{1+R}(1-p^*).$$

可以解出 $p^*=\frac{b-(1+R)}{b-a}$，所以，为了使 $0<p^*<1$，还必须满足 $a<1+R<b$.

（2）二叉模型下的欧式未定权益的定价

处理 Black-Scholes 模型的经验告诉我们，可以假定未定权益 $f(S_N)$在时刻 $n(<N)$的价格为 $V_n \overset{\text{def}}{=} V(n,S_n)$. 我们沿用套期的方法计算它. 设在时刻 n 卖空此未定权益，并待定买进 Δ_n 份标的证券，并将此投资组合的价值记为 X_n，则

$$X_n = \Delta_n S_n - V_n.$$

在 S_n 已知的条件下，要用这个投资组合来抵消在时刻 $n+1$ 由$\frac{S_{n+1}}{S_n}$带来的随机性. 由自融资假定，这个投资组合在时刻 $n+1$ 的价格应为

$$X_{n+1} = \Delta_n S_{n+1} - V_{n+1} = \Delta_n S_n \frac{S_{n+1}}{S_n} - V\left(n+1, S_n \frac{S_{n+1}}{S_n}\right). \tag{9.36}$$

为了使 X_{n+1}不再存在随机性，就是要使它的取值不再依赖于 a,b，也就是

$$\Delta_n S_n a - V(n+1, aS_n) = \Delta_n S_n b - V(n+1, bS_n).$$

为了保证这个等式成立，我们只需把待定的 Δ_n 取为

$$\Delta_n = \frac{V(n+1, bS_n) - V(n+1, aS_n)}{(b-a)S_n}. \tag{9.37}$$

利用无套利假定，这个消失了随机性的投资组合的时间发展应该按银行利率进行，即

$$X_{n+1} = (1+R)X_n \tag{9.38}$$

(如果取不等号就会出现如 Black-Scholes 模型的推导中那样的套利). 于是有

$$\begin{aligned} X_n &= \frac{1}{1+R}X_{n+1} = \frac{1}{1+R}[\Delta_n S_n a - V(n+1, aS_n)] \\ &= \frac{1}{1+R}\left\{\frac{a}{b-a}[V(n+1, bS_n) - V(n+1, aS_n)] - V(n+1, aS_n)\right\} \\ &= \frac{1}{1+R}\frac{1}{b-a}[aV(n+1, bS_n) - bV(n+1, aS_n)]. \end{aligned}$$

利用(9.37)式和(9.38)式得到

$$\begin{aligned} V(n, S_n) &= \Delta_n S_n - X_n \\ &= \frac{V(n+1, bS_n) - V(n+1, aS_n)}{b-a} - \frac{1}{1+R}\frac{1}{b-a}[aV(n+1, bS_n) - bV(n+1, aS_n)] \\ &= \frac{1}{1+R}\frac{1}{b-a}[(1+R-a)V(n+1, bS_n) + (b-(1+R))V(n+1, aS_n)] \\ &= \frac{1}{1+R}[(1-p^*)V(n+1, bS_{n+1}) + p^* V(n+1, aS_{n+1})]. \end{aligned}$$

即函数 $V(n, x)$ 满足后向差分方程

$$V(n, x) = \frac{1}{1+R}[(1-p^*)V(n+1, bx) + p^* V(n+1, ax)]. \tag{9.39}$$

并且由 $V(N, S_N) = f(S_N)$ 得到终端条件

$$V(N, x) = f(x). \tag{9.40}$$

用数学归纳法便可得到方程(9.39)满足条件(9.40)的解为

$$V(n, x) = \frac{1}{(1+R)^{N-n}}\sum_{k=0}^{N-n} C_{N-n}^k p^{*k}(1-p^*)^{(N-n)-k} f(a^k b^{(N-n)-k} x) \quad (n < N), \tag{9.41}$$

于是未定权益的贴水为

$$V(0, S_0) = \frac{1}{(1+R)^N}\sum_{k=0}^{N} C_N^k p^{*k}(1-p^*)^{N-k} f(a^k b^{N-k} S_0).$$

在实际计算中，可以用二叉树的后向递推法使用上述公式，通过树上操作反向地逐步求出欧式未定权益在各个时刻的价格.

9.3.2 Black-Scholes 模型的二叉模型近似

假设风险证券的价格 $S_t(t\leqslant T)$由 Black-Scholes 模型描述. 将$[0,T]$作 m 等分. 令

$$S_n^{(m)} \stackrel{\text{def}}{=} S_{\frac{nT}{m}}, \quad \widetilde{S}_n^{(m)} \stackrel{\text{def}}{=} \frac{1}{(1+R)^n} S_n^{(m)}. \tag{9.42}$$

这时要求$\widetilde{S}_T=\widetilde{S}_m^{(m)}$, 即要求 $\mathrm{e}^{-rT}S_T=\dfrac{1}{(1+R)^m}S_T$, 从而要求 $\mathrm{e}^{-rT}=\dfrac{1}{(1+R)^m}$, 由此可由 r, m 求出对应的 R. 记

$$\widetilde{T}_n \stackrel{\text{def}}{=} \frac{\widetilde{S}_n^{(m)}}{\widetilde{S}_{n-1}^{(m)}} = \frac{1}{1+R}\frac{S_n^{(m)}}{S_{n-1}^{(m)}}. \tag{9.43}$$

再假定$\dfrac{S_n^{(m)}}{S_{n-1}^{(m)}}$服从风险中性的概率分布$\begin{pmatrix} a & b \\ p^* & 1-p^* \end{pmatrix}$$\left(其中 p^*=\dfrac{b-(1+R)}{b-a}\right)$. 于是$\{\log \widetilde{T}_n\}$独立同分布, 而且

$$\log \widetilde{T}_n \sim \begin{pmatrix} \log \dfrac{a}{1+R} & \log \dfrac{b}{1+R} \\ p^* & 1-p^* \end{pmatrix}.$$

我们取

$$\log \frac{a}{1+R} = -\sigma\sqrt{\frac{T}{m}}, \quad \log \frac{b}{1+R} = \sigma\sqrt{\frac{T}{m}}. \tag{9.44}$$

当 m 很大时有

$$\begin{aligned} p^* &= \frac{b-(1+R)}{b-a} = \frac{\dfrac{b}{1+R}-1}{\dfrac{b}{1+R}-\dfrac{a}{1+R}} = \frac{\mathrm{e}^{\sigma\sqrt{\frac{T}{m}}}-1}{\mathrm{e}^{\sigma\sqrt{\frac{T}{m}}}-\mathrm{e}^{-\sigma\sqrt{\frac{T}{m}}}} \\ &\approx \frac{\sigma\sqrt{\dfrac{T}{m}}+\dfrac{1}{2}\sigma^2\dfrac{T}{m}}{2\sigma\sqrt{\dfrac{T}{m}}} = \frac{1}{2}+\frac{1}{4}\sigma\sqrt{\frac{T}{m}}. \end{aligned} \tag{9.45}$$

于是

$$\begin{aligned} E(\log \widetilde{T}_n) &= p^* \log \frac{a}{1+R} + (1-p^*)\log \frac{b}{1+R} \\ &\approx \left(\frac{1}{2}+\frac{1}{4}\sigma\sqrt{\frac{T}{m}}\right)\left(-\sigma\sqrt{\frac{T}{m}}\right)+\left(\frac{1}{2}-\frac{1}{4}\sigma\sqrt{\frac{T}{m}}\right)\left(\sigma\sqrt{\frac{T}{m}}\right) \\ &= -\frac{\sigma^2}{2}\frac{T}{m} = o(1) \quad (m \to +\infty). \end{aligned}$$

从而

$$\mathrm{var}(\log \widetilde{T}_n) \approx E(\log \widetilde{T}_n)^2 = p^*\left(\log \frac{a}{1+R}\right)^2 + (1-p^*)\left(\log \frac{b}{1+R}\right)^2$$
$$\approx \left(\frac{1}{2}+\frac{1}{4}\sigma\sqrt{\frac{T}{m}}\right)\sigma^2\frac{T}{m} + \left(\frac{1}{2}-\frac{1}{4}\sigma\sqrt{\frac{T}{m}}\right)\sigma^2\frac{T}{m} = \sigma^2\frac{T}{m}.$$

进而,对于 $t\leqslant T$ 由中心极限定理,当 m 大时,$\log \widetilde{S}^{(m)}_{\left[\frac{t}{T/m}\right]} = \sum\limits_{n=0}^{\left[\frac{t}{T/m}\right]} \log \widetilde{T}_n$ 的分布应与正态分布 $N\left(-\frac{\sigma^2}{2}\frac{T}{m}\left[\frac{t}{T/m}\right], \sigma^2\frac{T}{m}\left[\frac{t}{T/m}\right]\right)$近似,即与正态分布 $N\left(-\frac{\sigma^2}{2}t, \sigma^2 t\right)$近似.

用更精细的讨论(用多维的中心极限定理)可以得到:随机过程 $\log \widetilde{S}^{(m)}_{\left[\frac{t}{T/m}\right]}$ $(t\leqslant T)$的有限维分布与漂移 Brown 运动 $\sigma B_t - \frac{\sigma^2}{2}t$ 有相同的有限维分布. 也就是说,在 p^* 所对应的概率 $P^*(\cdot)$下,证券的折现价格$\widetilde{S}^{(m)}_{\left[\frac{t}{T/m}\right]}$ $(t\leqslant T)$作为随机过程的有限维分布与几何 Brown 运动 $\mathrm{e}^{\sigma B_t - \frac{\sigma^2}{2}t}$的有限维分布相同. 而后者正是风险中性的 Black-Scholes 模型. 这说明了离散二叉模型正是 Black-Scholes 模型的近似.

总之,对于给定的波动率 σ 和利率 r 及充分大的 m,可以求出对应的离散利率 R,二叉模型的参数 a,b 和公平概率 p^*,就可以用二叉模型$\widetilde{S}^{(m)}_{\left[\frac{t}{T/m}\right]}$来近似 Black-Scholes 模型. 而前者的计算是很简单的,可以用二叉树的后向传递法,通过树上操作求出未定权益在各个时刻的近似价格.

9.3.3 二叉模型的美式未定权益概要(美式权益$\{f(S_n): n\leqslant N\}$的定价,套期与消费过程)

(1) 离散模型的美式未定权益

我们不妨假定标的证券在开始时刻的价格 S_0 是常数.

定义 9.8 如果容许证券的未定权益的持有人在$[1,N]$中的任意时刻 n 都可以执行(但不在时刻 0 执行),将持有人在时刻 n 执行此权益的所得记为 X_n,它是一个随机变量,一般是 $S_n,\cdots,S_1$ 的一个函数(故 X_n 是(S_n)可知的). 与这种金融证券相系的未定权益序列$\{X_n: n\leqslant N\}$称为**美式未定权益**.

注意欧式未定权益就是 X_N,只是一个随机变量,而美式未定权益是一个随机变量序列. 又因为美式未定权益可以在$[1,N]$中的任意时刻执行,所以它的定价显然应该比欧式未定权益 X_N 的定价要高. 在本节中,我们将对二叉模型,给出美式未定权益的定价. 而对于 Black-Scholes 模型,则可以与欧式未定权益一样,用二叉模型得到近似.

如果 $X_n=(S_n-K)^+$,则此未定权益称为**美式看涨期权**,同样地,如果 $X_n=(K-S_n)^+$,则此未定权益称为**美式看跌期权**.

对二叉模型，$\frac{S_{n+1}}{S_n}$在风险中性的概率 P^* 下的分布为

$$\begin{pmatrix} a & b \\ p^* & q^* \end{pmatrix}, \quad p^* = \frac{b-(1+R)}{b-a}, \quad q^* = 1-p^*. \tag{9.46}$$

记 S_n 的折现为

$$\widetilde{S}_n = \frac{S_n}{(1+R)^n},$$

其中 R 为单位时间的银行利率. 我们将美式未定权益$\{X_n, n\leqslant N\}$在时刻 n 的定价记为 U_n（它也是随机变量）. 显见，在时刻 N 应有

$$U_N = X_N. \tag{9.47}$$

而$\{X_n, n\leqslant N\}$在时刻 $N-1$ 的定价 U_{N-1}，则决定于当时执行的收益与在下一个时刻执行的预期收益中的较大者. 由于它在下一时刻执行的预期收益应该是在时刻 N 到期的欧式未定权益在时刻 $N-1$ 的定价的折现，为 $E^*\left(\frac{U_N}{(1+R)^N}\middle| S_{N-1},\cdots,S_1\right)$. 于是

$$U_{N-1} = \max\left\{X_{N-1}, (1+R)^{N-1}E^*\left(\frac{U_N}{(1+R)^N}\middle| S_{N-1},\cdots,S_1\right)\right\}. \tag{9.48}$$

又因为 $U_N = X_N$，上式就是

$$U_{N-1} = \max\left\{X_{N-1}, (1+R)^{N-1}E^*\left(\frac{X_N}{(1+R)^N}\middle| S_{N-1},\cdots,S_1\right)\right\}. \tag{9.48$'$}$$

它可以改写为

$$U_{N-1} = (1+R)^{N-1}\max_{m\geqslant N-1} E^*\left(\frac{X_m}{(1+R)^m}\middle| S_{m-1},\cdots,S_1\right). \tag{9.48$''$}$$

一般地，美式未定权益在时刻 n 的定价，应该决定于当时执行的收益与把下一个时刻定价 U_{n+1}作为时刻 $n+1$ 到期的欧式未定权益的定价中的较大者. 由于在下一时刻到期的欧式未定权益 U_{n+1}的定价的折现应为 $E^*\left(\frac{U_{n+1}}{(1+R)^{n+1}}\middle| S_n,\cdots,S_1\right)$. 于是有

$$U_n = \max\left\{X_n, (1+R)^n E^*\left(\frac{U_{n+1}}{(1+R)^{n+1}}\middle| S_n,\cdots,S_1\right)\right\}. \tag{9.49}$$

由于 $U_n\geqslant X_n$，用条件期望的性质及归纳法不难证明

$$U_n = (1+R)^n \max_{m\geqslant n}\left(E^*\left(\frac{X_m}{(1+R)^m}\middle| S_n,\cdots,S_1\right)\right), \tag{9.50}$$

$$U_1 = \max_{m\geqslant 0} E^*\left(\frac{X_m}{(1+R)^m}\right). \tag{9.51}$$

于是我们得到下面的结论.

命题 9.8 美式未定权益的折现定价$\widetilde{U}_n\left(\widetilde{U}_n = \frac{U_n}{(1+R)^n}, n\leqslant N\right)$是$(S_n)$上鞅列，而且

它是满足 $U_n \geqslant X_n$ 条件的上鞅列中的最小的一个，所以 $\widetilde{U}_n$ 称为**控制 $\widetilde{X}_n$ 的最小上鞅列**，也称为 $\widetilde{X}_n$ 的 **Snell 包络**，其中 $\widetilde{X}_n = \dfrac{X_n}{(1+R)^n}$ 为未定权益的折现.

证明　前一个结论直接得自(9.49)式. 又若 Y_n 是另一个控制 $\widetilde{X}_n$ 的上鞅列，我们用后向归纳法来证明 $Y_n \geqslant \widetilde{U}_n$. 对于 $n=N$，我们有 $Y_N \geqslant \widetilde{X}_N = \widetilde{U}_N$. 假定 $n=m$ 时有 $Y_m \geqslant \widetilde{U}_m$，那么

$$Y_{m-1} \geqslant E^*(Y_m \mid S_{m-1}, \cdots, S_1) \geqslant E^*(\widetilde{U}_m \mid S_{m-1}, \cdots, S_1).$$

再利用(9.49)式便得 $Y_{m-1} \geqslant \widetilde{U}_{m-1}$. 由归纳法推出结论成立.

更一般地，对于任意一个取值于 $[n,N]$ 的 (S_k) 停时 τ，对它的不同可能取值用全期望公式还可以得到

$$U_n = (1+R)^n \max_{\text{一切停时}\tau\in[n,N]} E^*\left[\frac{X_\tau}{(1+R)^\tau}\,\middle|\, S_n, \cdots, S_1\right]. \tag{9.50$'$}$$

所以，美式未定权益在初始时刻的贴水又可以表示为

$$U_1 = \max_{\text{一切停时}\tau\in[0,N]} E^*\left[\frac{X_\tau}{(1+R)^\tau}\right]. \tag{9.51$'$}$$

显然，(9.50)′式与(9.51)′式并不能用于实算. 但是，我们可以用它来讨论未定权益的持有人在理论上于什么时刻执行最有利，这种有利时刻就是**最佳执行时刻**，它是一种最佳停时.

定义 9.9(未定权益的最佳执行时刻)　任意一个 (S_k) 停时 τ^*，只要满足

$$U_1 = E^*\left[\frac{X_{\tau^*}}{(1+R)^{\tau^*}}\right]\left(= \max_{\text{一切停时}\tau\in[0,N]} E^*\left[\frac{X_\tau}{(1+R)^\tau}\right]\right),$$

就称 τ^* 为**美式未定权益的一个最佳执行时刻**，其中最后一个等号正说明了它的最佳性质. 由这个定义可以看到，最佳执行时刻并不一定唯一.

在一般时刻 n，(9.50)′式指出了定价 $U_n \geqslant X_n$(该时刻执行的所得权益). 所以在直观上对未定权益的持有人不论在什么有利的时刻执行. 所得的收益也不会超过定价. 但是在定价随机变量与执行所得的权益随机变量相等的那些时刻(一般都是随机时刻)执行，所得的收益就达到了它的上界. 因此，这些随机时刻就应是最佳执行时刻. 于是我们有下面的定理.

定理 9.9　将定价与未定权益首次相等的时刻(是一个停时)记为 τ_0，即

$$\tau_0 = \min\{n > 0 : U_n = X_n\}.$$

那么，它也是美式未定权益的一个最佳执行时刻.

此定理在直观上是显然的，但是其证明则需要用鞅列在停时上的停止定理，本书从略.

注 定价的折现 $\widetilde{U}_n$ 是(S_n)上鞅列，这就是说，由预期美式未定权益的收益而得到的定价平均地是下降的. 它的直观意义很清楚，时间越早对持有人有更大的选择余地，定价当然就应该高一些. 所以，美式未定权益的定价作为随机变量(它按上鞅的规律发展)，其平均值是时间的下降序列. 这一点正体现了定价在平均的意义上，对于买方与卖方的公平性.

另一方面，即使持有人在最佳执行时刻执行，所得到收益的折现也不过就是当时的定价. 那么，购买这种未定权益能得到什么好处呢？事实上，虽然定价在平均意义上不过是一种公平的博采，但是正如许多公平的博采一样，人们还是抱着试试运气的心理参加投资，希望在个别的情形中，得到的报酬会远高于平均所得. 显然，在另外的一些情形，投资人也可能得到远低于平均的报酬. 那么，为什么这种金融工具会广泛流行呢？事实上，对于全社会而言，这样的金融工具正是利用人们普遍的趋利性的一种融资的手段，做到既筹集了资金，又分散了社会的金融风险，促进资本有效地流动. 当然，这并不包括黑户操纵与违规操作. 对于投资人来说，在健全的金融体制下，虽然有人会得利，也有人会受到损失，只要局限在经济力量容许的范围内，利用闲散资金投资，也是对金融领域的某种参与，况且在正常的情形，标的证券还可以得到分红. 然而，任何金融创新都是一把双刃剑，在起正面作用的时候，又常常被利用作为投机手段，这可能带来另一种新的风险. 这是人类社会的普遍的客观规律.

需要特别注意的是，实际市场的炒作与风险，远比数学模型复杂. 数学模型只是平均行为的一个非常粗放的参考，哪种模型近似得较为合理一些，应该以实证数据为依据.

(2) 二叉模型美式未定权益$\{f(S_n), n\leqslant N\}$的定价

本段我们讨论二叉模型美式未定权益的一种常见的特殊情形，即存在 $f(x)$使对于任意 $n\leqslant N$，都有

$$X_n = f(S_n)$$

的情形(典型的例子是美式看涨期权和美式看跌期权，它们分别对应于 $f(x)=(x-K)^+$ 和 $f(x)=(K-x)^+$)，此时，由(9.47)式～(9.49)式可以直接推出

$$U_N = f(S_N) \stackrel{\text{def}}{=} U_N(S_N) \quad (\text{其中 } U_N(x) = f(x)), \tag{9.52}$$

$$\begin{aligned} U_{N-1} &= \max\left\{f(S_{N-1}), \frac{1}{1+R}E^*(f(S_N) \mid S_{N-1},\cdots,S_1)\right\} \\ &= \max\left\{f(S_{N-1}), \frac{1}{1+R}[p^* f(aS_{N-1}) + q^* f(bS_{N-1})]\right\} \stackrel{\text{def}}{=} U_{N-1}(S_{N-1}), \end{aligned}$$

其中

$$U_{N-1}(x) = \max\left\{f(x), \frac{1}{1+R}[p^* f(ax) + q^* f(bx)]\right\}.$$

对于一般的 $n<N$，用后向归纳法可得

$$U_n = \max\left\{f(S_n), \frac{1}{1+R}[p^* U_{n+1}(aS_n) + q^* U_{n+1}(bS_n)]\right\} \stackrel{\text{def}}{=} U_n(S_n),$$

其中

$$U_n(x) = \max\left\{f(x), \frac{1}{1+R}[p^* U_{n+1}(ax) + q^* U_{n+1}(bx)]\right\}. \tag{9.53}$$

由此我们归纳出下面的定理.

定理 9.10 只要后向地逐个从以下的终端已知的极值函数方程组：

$$\begin{cases} U_N(x) = f(x), \\ U_n(x) = \max\left\{f(x), \frac{1}{1+R}[p^* U_{n+1}(ax) + q^* U_{n+1}(bx)]\right\} \quad (n < N) \end{cases} \tag{9.54}$$

解出 $\{U_n(x): n \leqslant N\}$，便得到美式未定权益的定价 $U_n = U_n(S_n)$.

定义 9.10(美式未定权益的定价函数组) 由定理 9.10 所确定的 $\{U_n(x): n \leqslant N\}$，称为美式未定权益 $\{f(S_n)\}(n \leqslant N)$ 的定价函数组.

于是对二叉模型，美式未定权益的定价可以由极值函数方程组的数值求解得到. 即为了达到最佳执行，只须用下述算法：

① 对于 $f(x)$，计算出定价函数组 $\{U_n(x): n \leqslant N\}$；

② 对于证券的实时市场价 $S_0, S_1, \cdots$，实时地计算 $f(S_1), f(S_2), \cdots$ 与 $U_1(S_1)$, $U_2(S_2), \cdots$，并实时地依次比较对应的值；

③ 在计算(2)的过程中，一旦发现它们在某个时刻相等，则此时刻就是最佳执行时刻. 即：若样本列 $f(S_1), f(S_2), \cdots$ 与定价列 $U_1(S_1), U_2(S_2), \cdots$ 第一次相等出现于 m，则未定权益的持有人就在此时刻 m 执行.

(3) 美式未定权益的套期与消费过程

由于折现价格序列 $\widetilde{U}_n$ 是上鞅列，于是

$$\widetilde{C}_n = \widetilde{U}_n - E^*(\widetilde{U}_{n+1} \mid S_n, \cdots, S_1) \geqslant 0. \tag{9.55}$$

即随机变量列 $\widetilde{C}_n$ 为非负的. 又因为美式未定权益在任意时刻都可以执行，我们只能仿照欧式未定权益的套期公式来套期. 也就是说，在时刻 n 的套期标的证券的数目取为

$$\Delta_n = \frac{U_{n+1}(bS_n) - U_{n+1}(aS_n)}{(b-a)S_n}. \tag{9.56}$$

但是与欧式未定权益的情形唯一的不同，就是用美式未定权益的定价 $\{U_n\}$ 代替了欧式未定权益的定价 $\{V_n\}$. 令 $C_n = (1+R)^n \widetilde{C}_n$，于是

$$C_n = U_n - \frac{1}{1+R} E^*(U_{n+1} \mid S_n, \cdots, S_1). \tag{9.57}$$

由于

$$E^*(U_{n+1} \mid S_n, \cdots, S_1) = p^* U_{n+1}(aS_n) + q^* U_{n+1}(bS_n),$$

有

$$\begin{aligned} U_{n+1}(aS_n) &= q^*[U_{n+1}(aS_n) - U_{n+1}(bS_n)] + p^* U_{n+1}(aS_n) + q^* U_{n+1}(bS_n) \\ &= -[(1+R) - a]\Delta_n S_n + E^*(U_{n+1} \mid S_n, \cdots, S_1) \\ &= -(1+R)\Delta_n S_n + \Delta_n a S_n - (1+R)(C_n - U_n). \end{aligned}$$

完全类似地还可以证明

$$U_{n+1}(bS_n) = -(1+R)\Delta_n S_n + \Delta_n b S_n - (1+R)(C_n - U_n).$$

将 $S_{n+1}=aS_n$ 与 $S_{n+1}=bS_n$ 两种可能的情形合起来,就得到

$$U_{n+1} = \Delta_n S_{n+1} + (1+R)(U_n - C_n - \Delta_n S_n).$$

将它改写为如下等式

$$U_{n+1} - \Delta_n S_{n+1} = (1+R)(U_n - \Delta_n S_n) - (1+R)C_n. \tag{9.58}$$

这公式具有比较清楚的金融含义,它说明在美式未定权益情形,即使用合适的套期 Δ_n,也不能使余下的资产 $U_n - \Delta_n S_n$ 为无风险资产.只有在时刻 n 的套期余额中消耗了 C_n 以后的资产,到下一个时刻 $n+1$ 的价值才是时刻 $n+1$ 的套期后的余额.这个随机序列$\{C_n\}$称为**美式未定权益的消费过程**.

一般地,不管是什么模型,美式未定权益的套期总需要有一个消费过程去补充.在理论上知道美式未定权益的定价需要补充一个消费过程,是对美式未定权益定价模型认识的一个实质性的进展.

9.4 随机利率与债券利率的期限结构

9.4.1 s-零息债券

定义 9.11 不付红利,而且在时刻 s 到期的债券称为 **s-零息债券**(zero-coupon).

s-零息债券(券面价值是指到期时的价值)进入市场后,在市场操作下,按某个随机利率增值.记时刻 $t(<s)$的 1 元钱按市场发展在债券到期时刻 s 的价值("本利和")为 $R(t,s)$.将时刻 s 的 1 元钱在时刻 $t(<s)$的定价记为 $P(t,s)$.显见

$$P(t,s) = \frac{1}{R(t,s)}, \quad P(s,s) = 1. \tag{9.59}$$

一般地,$P(t,s)$是随机过程.请注意,由于 s-零息债券与 t-零息债券在市场的表现是不一样的,所以即使在无套利的假定下,一般地

$$P(0,s) \neq P(0,t)P(t,s) \quad (t < s). \tag{9.60}$$

这是因为(9.60)式的左边是由初始时刻的市场信息所确定,而右边是由直至时刻 t 为止的信息所确定的,因此一般不会相等.

但是如果定价$\{P(t,s), 0 \leqslant t \leqslant s\}$不是随机过程而是确定性函数时,对于债券而言,在

时刻 t 与在初始时刻有相同的信息. 此时在无套利假定下就有

$$P(0,s)=P(0,t)P(t,s)\quad(t<s).$$

9.4.2 零息债券导出的各种随机利率概念

定义 9.12(平均远期利率 (到期收益率)) 记

$$\bar{f}(t,s)=-\frac{\log P(t,s)}{s-t}\quad(t<s),\tag{9.61}$$

即

$$P(t,s)=\mathrm{e}^{-\bar{f}(t,s)(s-t)}\quad(t<s).\tag{9.61$'$}$$

$\bar{f}(t,s)$称为 s-零息债券的**平均远期利率**. 这里的"远"的含义是指参考时间 $t>0$,即参考时间不是初始时刻的利率,它是在将来的时刻 $t(<s)$展望 s-债券的平均利率. 由 $1=P(t,s)\mathrm{e}^{\bar{f}(t,s)(s-t)}(t<s)$,说明平均远期利率就是:在时刻 t 将 s-债券换为比它早到期的 t-债券所需付的利率.

在时刻 $t(<s)$固定时,平均远期利率作为 $s-t$ 的函数的图形变化,称为**利率的期限结构**.

更一般地有下面的定义.

定义 9.13 对于涉及不同的到期时刻的两种零息债券,定义

$$\bar{f}(t,s_1,s_2)=-\frac{\log P(t,s_2)-\log P(t,s_1)}{s_2-s_1}\quad(0<t<s_1<s_2),\tag{9.62}$$

即

$$P(t,s_1)=P(t,s_2)\mathrm{e}^{\bar{f}(t,s_1,s_2)(s_2-s_1)}.$$

这说明$\bar{f}(t,s_1,s_2)$就是在时刻 t 将 s_2-债券换为比它早到期的 s_1-债券所需付的利率.

按定义有

$$\bar{f}(t,s)=\bar{f}(t,t,s).\tag{9.63}$$

定义 9.14(平均即期利率,spot rate) $\bar{f}(0,s)$称为**平均即期利率**.

于是我们有 $P(0,s)=P(0,0)\mathrm{e}^{-s\bar{f}(0,s)}$,此即(9.61)$'$式的特殊情形.

定义 9.15(远期利率,instantaneous forward rate) 若 $P(t,s)$对 s 可微分,则

$$f(t,s)\xlongequal{\text{def}}-\frac{\partial\log P(t,s)}{\partial s}\Big(=\lim_{\Delta s\to 0}\bar{f}(t,s,s+\Delta s)\Big)\tag{9.64}$$

称为**远期利率**. 于是有

$$P(t,s)=\mathrm{e}^{-\int_t^s f(t,u)\mathrm{d}u},\tag{9.65}$$

$$\bar{f}(t,s)=\frac{1}{s-t}\int_t^s f(t,u)\mathrm{d}u.\tag{9.66}$$

(9.66)式也可由$\frac{\mathrm{d}}{\mathrm{d}s}[(s-t)\overline{f}(t,s)]=-\frac{\mathrm{d}\log P(t,s)}{\mathrm{d}s}=f(t,s)$证得. 由(9.65)式与(9.66)式可以看出，给出价格随机过程$\{P(t,s),0\leqslant t\leqslant s\}$与给出远期利率随机过程$\{f(t,s),0\leqslant t\leqslant s\}$是等效的.

定义 9.16(短期利率，short rate)　$r(t)\overset{\mathrm{def}}{=}f(t,t)$称为**短期利率**.

短期利率就是普通习惯上所称的利率，不过它是随机的. 我们有

$$r(t)=f(t,t)=\overline{f}(t,t). \tag{9.67}$$

证明　$f(t,t)$与$\overline{f}(t,t)$按定义都是$\lim\limits_{s\downarrow t}-\frac{\log P(t,t+(s-t))}{s-t}$.

例 9.1(非随机情形)　这时$\{P(t,s),0\leqslant t\leqslant s\}$是确定性函数，因而所涉及的各种利率也都是确定性的函数. 在不十分剧烈炒作的场合下，可以用此看法对实际市场作近似. 此时各种利率可以有如下的一些换算公式：对于$0\leqslant t\leqslant s$，有

(1) $P(0,s)=P(0,t)P(t,s)\quad(t<s)$.

(2) $f(0,s)=f(t,s)$，即$f(t,s)$不依赖t. 从而$f(t,s)=f(s,s)=r(s)$.

(3) $\overline{f}(t,s)=\frac{1}{s-t}\int_t^s r(u)\mathrm{d}u$，从而有$P(t,s)=\mathrm{e}^{-\int_t^s r(u)\mathrm{d}u}$.

(4) $\overline{f}(t,s)=\frac{s\overline{f}(0,s)-t\overline{f}(0,t)}{s-t}$.　　(9.68)

一般从交易所的挂牌数据，可以得到的平均即期利率$\overline{f}(0,t)$，$\overline{f}(0,s)$，再用(9.68)式算出平均远期利率$\overline{f}(t,s)$. 而在实际交易中，人们用的正是平均远期利率. 由(9.68)式还可以得到

$$\overline{f}(0,s)\geqslant\overline{f}(0,t)\Rightarrow\overline{f}(t,s)\geqslant\overline{f}(0,s)$$

(5) $r(t)=\overline{f}(0,t)+t\frac{\mathrm{d}}{\mathrm{d}t}\overline{f}(0,t)$.

证明　将(4)改写为$\overline{f}(t,s)=\overline{f}(0,s)+(\overline{f}(0,s)-\overline{f}(0,t))\frac{t}{s-t}$. 令$s\downarrow t$，便得(5).

9.4.3　资产定价基本定理与利率衍生证券

假设时刻t的短期利率$r(t)$可由市场在t以前的信息(例如某个 Brown 运动B_t^*在t以前的信息$(B_t^*:u\leqslant t)$)所完全确定. 此时我们有下面的定理.

定理 9.11(资产定价基本定理)　市场无套利性等价于具有风险中性概率P^*，在此概率下的折现值$\widetilde{P}_t\overset{\mathrm{def}}{=}\frac{P(t,s)}{\mathrm{e}^{\int_0^t r(u)\mathrm{d}u}}$为$(B_t^*)$鞅，其中$P(t,s)$是1元的$s$-零息债券在时刻$t(<s)$的价格，故而

$$P(t,s)=\mathrm{e}^{\int_0^t r(u)\,\mathrm{d}u}E^*(\widetilde{P}_s \mid B_u^*: u\leqslant t).$$

即 s-零息债券的无套利定价为

$$P(t,s)=E^*(\mathrm{e}^{-\int_t^s r(u)\,\mathrm{d}u} \mid B_u^*: u\leqslant t). \tag{9.69}$$

另一方面,(9.69)式也保证了 $\widetilde{P}_t \stackrel{\text{def}}{=} \dfrac{P(t,s)}{\mathrm{e}^{\int_0^t r(u)\,\mathrm{d}u}}$ 是(B_t^*)鞅.

由此看出,只要给定了短期利率 $r(t)$的风险中性模型,就可以由(9.69)式得到 s-零息债券的无套利定价.

定义 9.17 一个金融合约的未定权益,如果依赖于未来的利率,或债券的价格,则此合约就称为**利率未定权益**,或**利率衍生证券**.

随机利率模型的研究,是要给出 s-债券的定价,远期利率及利率未定权益的定价.

9.4.4 利率的风险中性模型(远期利率的 HJM 模型,短期利率模型)

一般地,只要给定短期利率 $r(t)$,由(9.69)式定义的 $P(t,s)$就保证了$\widetilde{P}_t \stackrel{\text{def}}{=} \dfrac{P(t,s)}{\mathrm{e}^{\int_0^t r(u)\,\mathrm{d}u}}$ 是(B_t^*)鞅,从而 $P(t,s)$就是无套利定价,此时用来确定 $r(t)$的随机微分方程的模型,就是风险中性模型. 所以短期利率的风险中性模型可以任意给定,不带什么约束条件(事实上,从(9.69)式求出的远期利率为 $f(t,s)=E^*(r(s)\mathrm{e}^{-\int_t^s r(u)\,\mathrm{d}u} \mid B_u^*: u\leqslant t)$,得到的 $f(t,t)$ 正好等于 $r(t)$). 另一方面,如果直接给出价格 $P(t,s)$满足随机微分方程,或者等价地,直接给出远期利率 $f(t,s)$满足的随机微分方程,由此确定的 $r(t)$就未必满足(9.69)式. 这时(9.69)式就成为约束条件. 只有满足了这个约束条件的 $P(t,s)$才是无套利定价,所对应的 $P(t,s)$或 $f(t,s)$的随机微分方程模型才是风险中性模型.

(1) 远期利率的 Heath-Jarrow-Morton 模型及其风险中性情形

由资产定价基本定理可以知道,为了使模型成为风险中性,充分必要条件是 $\widetilde{P}_t \stackrel{\text{def}}{=} \dfrac{P(t,s)}{\mathrm{e}^{\int_0^t r(u)\,\mathrm{d}u}}$ 是鞅. 由此利用(9.65)式及 Girsanov 定理,就能得到以下的定理.

定理 9.12(Heath-Jarrow-Morton 定理) 对于固定的 s,如果远期利率 $f(t,s)$是以下的 Ito 过程

$$\mathrm{d}f(t,s)=b(t,s)\mathrm{d}t+\sigma(t,s)\mathrm{d}B_t, \tag{9.70}$$

其中 $b(t,s)$,$\sigma(t,s)$在 s 固定下是(B_t)可知的随机过程. 那么,此模型是无套利(即风险中性的)的充要条件为:存在一个恒满足条件

$$E\mathrm{e}^{-\int_0^s \theta(t)\,\mathrm{d}B_t-\frac{1}{2}\int_0^s \theta(t)^2\,\mathrm{d}t}=1$$

的(B_t)可知的随机过程 $\theta(t)$,使

$$b(t,s)=\sigma(t,s)\left[\int_t^s \sigma(t,u)\,\mathrm{d}u+\theta(t)\right].$$

在条件成立时，由 Girsanov 定理可知，$B_t^* \stackrel{\text{def}}{=} B_t+\int_0^t \theta(u)\,\mathrm{d}u$ 也是 Brown 运动，且

$$\mathrm{d}f(t,s)=\sigma(t,s)\left[\int_t^s \sigma(t,u)\,\mathrm{d}u\right]\mathrm{d}t+\sigma(t,s)\,\mathrm{d}B_t^*. \tag{9.71}$$

再对 $P(t,s)=\mathrm{e}^{-\int_t^s f(t,u)\mathrm{d}u}$ 用 Ito 公式可得

$$\mathrm{d}P(t,s)=P(t,s)\left[r(t)\mathrm{d}t-\left(\int_t^s \sigma(t,u)\,\mathrm{d}u\right)\mathrm{d}B_t^*\right]. \tag{9.72}$$

可见在此情形下，s-零息债券的价格 $P(t,s)$ 的收益率为 $r(t)$，波动率为 $\int_t^s \sigma(t,u)\mathrm{d}u$（而 $\sigma(t,s)/f(t,s)$ 为远期利率的波动率）.

（2）短期利率风险中性模型

直接给出短期利率风险中性模型的有以下几种.

① **Vasicek 模型**

假定短期利率 $r(t)$ 满足如下的 Ornstein-Uhlenbeck 随机微分方程：

$$\mathrm{d}r(t)=a(b-r(t))\mathrm{d}t+\sigma\,\mathrm{d}B_t^*. \tag{9.73}$$

这个随机微分方程具有显式解：

$$r(t)=r(0)\mathrm{e}^{-at}+b(1-\mathrm{e}^{-at})+\sigma\mathrm{e}^{-at}\int_0^t \mathrm{e}^{as}\,\mathrm{d}B_s^*. \tag{9.74}$$

容易证明它是 Gauss 过程. 而作为 Gauss 过程的积分 $\int_0^t r(u)\mathrm{d}u$ 也是 Gauss 过程. 于是可以仿照风险中性概率推导 Black-Scholes 公式的计算过程算出(9.73)式中的条件期望. 经过仔细而并不复杂的推导计算，便可得到

$$P(t,s)=\mathrm{e}^{-A(t,s)-C(t,s)r(t)}, \tag{9.75}$$

其中

$$A(t,s)=\frac{1}{a^2}(ts-C(t,s))\left(a^2b-\frac{\sigma^2}{2}\right)+\frac{\sigma^2 C(t,s)^2}{4a},\quad C(t,s)=\frac{1-\mathrm{e}^{-a(s-t)}}{a}.$$

于是远期利率为

$$f(t,s)=\frac{C(t,s)r(t)+A(t,s)}{s-t},$$

它是短期利率的线性函数.

② **Hull-White 模型**

Hull-White 模型是 Vasicek 模型在时变系数情形的推广，即模型参数 a,b,σ 分别被推广为确定性的函数 $a(t),b(t),\sigma(t)$，即

$$\mathrm{d}r(t)=a(t)(b(t)-r(t))\mathrm{d}t+\sigma(t)\mathrm{d}B_t^*. \tag{9.74$'$}$$

与常数情形完全类似地可以算出定价过程仍是 $r(t)$ 的如下函数

$$P(t,s)=\mathrm{e}^{-A(t,s)-C(t,s)r(t)},\tag{9.75)'$$

只需要对 $A(t,s)$, $C(t,s)$ 作一些相应的变更.

(9.75)′式说明,此时 s-零息债券的价格 $P(t,s)(t\leqslant s)$ 是 $r(t)$ 的函数. 由于 $r(t)$ 是 Ito 过程. 用 Ito 公式就可得到

$$\mathrm{d}P(t,s)=P(t,s)[r(t)\mathrm{d}t-\sigma(t)C(t,s)\mathrm{d}B_t^*].\tag{9.76}$$

这就是说,$P(t,s)$ 的收益率是 $r(t)$,波动率是 $\sigma(t)C(t,s)$. (9.75)′式可以用于由实证数据校正模型的参变函数组 $a(t)$, $b(t)$, $\sigma(t)$. 但是,Hull-White 模型存在的缺点是,短期利率服从正态分布,因而可以取负值,这是不合理的.

Hull-White 模型下的债券看涨期权 对于时刻 $T(<s)$ 到期的,以 s-零息债券为标的证券的,且执行价格为 K 的欧式看涨期权,其未定权益为 $[P(T,s)-K]^+$. 于是它在时刻 t 的价格应该为

$$V_t\overset{\text{def}}{=\!=}E^*(\mathrm{e}^{-\int_t^s r(u)\mathrm{d}u}[P(T,s)-K]^+\mid B_u^*:u\leqslant t).\tag{9.77}$$

利用 $\left(r(t),\int_0^t r(u)\mathrm{d}u\right)$ 是 Gauss 过程,就可以计算出这种期权的贴水 V_0 的明显公式.

③ CIR(Cox-Ingersoll-Ross)模型

Cox-Ingersoll-Ross 假定短期利率 $r(t)$ 满足随机微分方程

$$\mathrm{d}r(t)=a(b-r(t))\mathrm{d}t+\sigma\sqrt{r(t)}\,\mathrm{d}B_t^*.\tag{9.78}$$

如果有 $d(\geqslant 2)$ 个彼此独立的 Brown 运动 $B_t^{(1)},\cdots,B_t^{(d)}$,又对于任意 $i\leqslant d$ 有

$$\mathrm{d}X_t^{(i)}=-\frac{a}{2}X_t^{(i)}\mathrm{d}t+\sigma\mathrm{d}B_t^{(i)}.$$

那么对于 $r(t)\overset{\text{def}}{=\!=}\sqrt{(X_t^{(1)})^2+\cdots+(X_t^{(d)})^2}$,可以证明存在 Brown 运动 B_t^* 使

$$\mathrm{d}r(t)=\left(\frac{\sigma^2 d}{4}-ar(t)\right)\mathrm{d}t+\sigma\sqrt{r(t)}\,\mathrm{d}B_t^*,$$

称为 Bessel 过程,所以 CIR 模型是 Bessel 过程. 因为 $r(t)$ 是扩散过程,求解不变密度的方程

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}\left(\frac{\sigma^2 r}{2}\varphi\right)-\frac{\mathrm{d}}{\mathrm{d}r}(a(b-r)\varphi)=0$$

就得到不变密度

$$\varphi(r)=Cr^{\frac{2ab-\sigma^2}{\sigma^2}}\mathrm{e}^{-\frac{2a}{\sigma^2}r}\quad(r>0),\tag{9.79}$$

其中

$$C=\left(\frac{2a}{\sigma^2}\right)^{\frac{2ab}{\sigma^2}}\frac{1}{\Gamma\left(\frac{2ab}{\sigma^2}\right)}.$$

CIR 模型的债券的定价 此时 $\{B_u^*:u\leqslant t\}$ 与 $\{r(u):u\leqslant t\}$ 具有相同的信息,所以

$$P(t,s)=E^*(\mathrm{e}^{-\int_t^s r(u)\mathrm{d}u}\mid B_u^*:u\leqslant t)=E^*(\mathrm{e}^{-\int_t^s r(u)\mathrm{d}u}\mid r(u):u\leqslant t).$$

由 $r(t)$ 的 Markov 性可知上式右边的量只依赖于 $r(t)$，即存在一个数值函数 $F(t,r,s)$ 使

$$P(t,s)=F(t,r(t),s).\tag{9.80}$$

假定 $F(t,r,s)$ 对 r 二阶连续可微，对 t 连续可微，那么 $\mathrm{e}^{\int_0^t r(u)\mathrm{d}u}F(t,r(t),s)$ 是一个 Ito 过程. 利用推广的 Ito 公式可算出

$$\begin{aligned}\mathrm{d}(\mathrm{e}^{\int_0^t r(u)\mathrm{d}u}F(t,r(t),s))&=\left[\frac{\partial F}{\partial t}+\frac{\sigma^2 r(t)}{2}\frac{\partial^2 F}{\partial r^2}+a(b-r(t))\frac{\partial F}{\partial r}-r(t)F\right]\mathrm{d}t\\&\quad+\sigma\sqrt{r(t)}\frac{\partial F}{\partial r}\mathrm{d}B_t^*.\end{aligned}$$

另一方面，由定义 $\mathrm{e}^{-\int_0^t r(u)\mathrm{d}u}F(t,r(t),s)=E^*(\mathrm{e}^{-\int_0^s r(u)\mathrm{d}u}\mid B_u^*:u\leqslant t)$ 是一个 (B_t^*) 鞅. 两者比较可知 F 必须满足偏微分方程：

$$\frac{\partial F}{\partial t}+\frac{\sigma^2 r}{2}\frac{\partial^2 F}{\partial r^2}+a(b-r)\frac{\partial F}{\partial r}-rF=0.\tag{9.81}$$

及终端条件

$$F(s,r,s)=1.\tag{9.82}$$

我们通过待定 $A(t,s)$，$C(t,s)$，探求方程的以下形式的解的可能性

$$F(t,r,s)=\mathrm{e}^{-A(t,s)-C(t,s)r}.\tag{9.83}$$

将(9.83)式代入(9.81)式，就得到 $A(t,s)$，$C(t,s)$ 关于 t 的二阶常微分方程. 求解得到

$$C(t,s)=\frac{\sinh(\gamma(s-t))}{\gamma\cosh(\gamma(s-t))+\frac{a}{2}\sinh(\gamma(s-t))},$$

$$A(t,s)=-\frac{2ab}{\sigma^2}\log\frac{\gamma\mathrm{e}^{\frac{a}{2}(s-t)}}{\gamma\cosh(\gamma(s-t))+\frac{a}{2}\sinh(\gamma(s-t))},$$

其中

$$\gamma=\frac{1}{2}\sqrt{a^2+2\sigma^2},$$

而 $\cosh(x)$，$\sinh(x)$ 分别是双曲余弦与双曲正弦：

$$\cosh(x)=\frac{\mathrm{e}^x+\mathrm{e}^{-x}}{2},\quad \sinh(x)=\frac{\mathrm{e}^x-\mathrm{e}^{-x}}{2}.$$

CIR 模型的债券看涨期权的定价　对于时刻 $T(<s)$ 到期的，以 s-零息债券为标的证券的，且执行价格为 K 的欧式看涨期权，其未定权益为 $[P(T,s)-K]^+$. 于是它在时刻 t 的价格应该为

$$\begin{aligned}V_t&\overset{\text{def}}{=\!=}E^*(\mathrm{e}^{-\int_t^s r(u)\mathrm{d}u}[P(T,s)-K]^+\mid B_u^*:u\leqslant t)\\&=E^*(\mathrm{e}^{-\int_t^s r(u)\mathrm{d}u}[F(T,r(T),s)-K]^+\mid B_u^*:u\leqslant t).\end{aligned}\tag{9.84}$$

与债券的定价类似地，可假定存在数值函数 $G(t,r,T,s)$ 使债券看涨期权的价格 $V_t=G(t,r(t),T,s)(t\leqslant T<s)$. 再利用 $\mathrm{e}^{-\int_0^t r(u)\mathrm{d}u}G(t,r(t),T,s)=E^*(\mathrm{e}^{-\int_0^s r(u)\mathrm{d}u}[P(T,s)-K]^+\mid B_u^*:u\leqslant t)$ 是一个(B_t^*)鞅，就得到 $G(t,r,T,s)$满足如下的偏微分方程

$$\frac{\partial G}{\partial t}+\frac{\sigma^2 r}{2}\frac{\partial^2 G}{\partial r^2}+a(b-r)\frac{\partial G}{\partial r}-rG=0 \quad (t\leqslant T) \tag{9.85}$$

及终端条件

$$G(T,r,T,s)=(F(T,r,s)-K)^+. \tag{9.86}$$

9.5 基于证券的随机利率的债券为币值单位折现的证券的未定权益的定价

设证券价格 S_t 与债券价格 b_t 分别满足随机微分方程

$$\mathrm{d}S_t=\mu(t,S_t)\mathrm{d}t+\sigma(t,S_t)\mathrm{d}B_t,$$
$$\mathrm{d}b_t=r(t,S_t)b_t\mathrm{d}t,$$

其中 B_t 为 Brown 运动，而债券 b_t 的市价是以依赖于证券值 S_t 的短期的连续随机利率 $r=r(t,S_t)$而增值的. 与 Black-Scholes 模型类似，将 $v(t,S_t)=\dfrac{\mu(t,S_t)-r(t,S_t)}{\sigma(t,S_t)}$称为证券 S_t 的**风险的市场价格**. 与 Black-Scholes 模型不同的是，这里不是用按非随机的利率增值的银行存款作为证券币制单位，而是按用与证券密切相关的随机利率增值的债券 b_t 作为该证券的币值单位.

于是，与 Black-Scholes 模型类似，如下确定的新的概率

$$P^*(A)=E(\mathrm{e}^{-\int_0^T v(t,S_t)\mathrm{d}B_s-\frac{1}{2}\int_0^T v(t,S_t)^2\mathrm{d}s}I_A)$$

称为**风险中性概率**. 在此概率下证券按债券的折现价格$\widetilde{S}_t\stackrel{\text{def}}{=}\dfrac{S_t}{b_t}$是$(S_t)$鞅.

对此模型的未定权益 $f(S_T)$在时刻 t 的定价为

$$F(t,S_t)=E^*\left[f(S_T)\frac{b_t}{b_T}\,\middle|\,S_t,\cdots,S_1\right].$$

同样，与 Black-Scholes 模型类似，价格函数 $F(t,x)$应满足如下的带终端条件的偏微分方程：

$$F_t+\frac{\sigma^2(t,x)}{2}x^2F_{xx}+r(t,x)xF_x-r(t,x)F=0,$$
$$F(T,x)=f(x).$$

习题 9

1. 对于风险中性随机利率按方程 $dr_t = \sigma dB_t + (a - br_t)dt$ 发展的金融债券，在时刻 T 的单位币值在时刻 $t(\leqslant T)$ 的折现为 $V(t,T) \stackrel{\text{def}}{=} E(e^{-\int_t^T r_s ds} \mid r_s: s \leqslant t)$. 证明 $\widetilde{V}(t,T) \stackrel{\text{def}}{=} V(t,T)e^{-\int_0^t r_s ds}$ 是 (r_t) 鞅，而且存在函数 $F(t,x;T)$，使 $V(t,T) = F(t,r_t;T)$. 又若 $F(t,x,T)$ 分别对 t,x 为一阶与二阶连续可微，证明它满足

$$F_t + \frac{\sigma^2}{2}F_{xx} + a(b-x)F_x - xF = 0,$$

$$F(T,x,T) = 1.$$

再令

$$F(t,x;T) = e^{A(t,T) - xB(t,T)},$$

并待定地确定

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a}, \quad A(t,T) = \left(b - \frac{\sigma^2}{2a^2}\right)\left[t - T + B(t) - \frac{\sigma^2}{4a}B^2(t)\right].$$

2. 对于随机利率按方程 $dr_t = \sigma dB_t + (a - br_t)dt$ 发展的金融债券，在时刻 T 的单位币值在时刻 $t(\leqslant T)$ 的折现为 $V(t,T) \stackrel{\text{def}}{=} E(e^{-\int_t^T r_s ds} \mid r_s: s \leqslant t)$. 证明 $\widetilde{V}(t,T) \stackrel{\text{def}}{=} V(t,T)e^{-\int_0^t r_s ds}$ 是 (r_t) 鞅.

3. 对于随机微分方程 $dr_t = \sigma\sqrt{r_t}\,dB_t + (a - br_t)dt$. 令 $G(t,x) \stackrel{\text{def}}{=} E_x(e^{-\lambda r_t - \mu\int_0^t r_u du})$，假定它分别对 t,x 为一阶与二阶连续可微，利用 $M_t \stackrel{\text{def}}{=} E(e^{-\lambda r_T - \mu\int_0^T r_u du} \mid r_u: u \leqslant t)$ 是鞅，证明

$$G_t = \frac{\sigma^2}{2}xG_{xx} + a(b-x)G_x - \mu xG = 0, \quad G(0,x) = e^{-\lambda x}.$$

第 10 章　Poisson 随机分析大意

10.1　Poisson 过程，非时齐的 Poisson 过程与复合 Poisson 过程的复习

本节内容可以参见一般的应用随机过程教科书. 本节的定理略去其证明.

10.1.1　Poisson 过程

定义 10.1　非负整数值随机过程称为计数过程.

定义 10.2　计数过程$\{N_t: t\geqslant 0\}$称为强度为 λ 的 Poisson 过程，如果$\{N_t: t\geqslant 0\}$是时齐的独立增量过程，满足 $N_0=0$，在有限时间区段内跳跃的次数是有限的，而且在小的时间区段 h 满足

$$P(N_{t+h}-N_t=1)=\lambda h+o(h),\quad P(N_{t+h}-N_t\geqslant 2)=o(h).$$

命题 10.1　非负整数值随机过程$\{N_t: t\geqslant 0\}$是 Poisson 过程等价于下述论述之一：

(1) $\{N_t: t\geqslant 0\}$是独立增量过程，满足 $N_0=0$，存在 $\lambda>0$，使对于任意 $s\geqslant 0$ 非负整数 k. 有

$$P(N_{s+t}-N_s=k)=\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}.$$

(2) $\{N_t: t\geqslant 0\}$是时齐的独立增量过程，满足 $N_0=0$，且有矩母函数

$$Ez^{zN_t}=\mathrm{e}^{\lambda(z-1)t}.$$

(3) 令 $\tau_k=\inf\{t: N_t=k\}$是第 k 个跳跃时刻，则 τ_k 是 k 个参数为 λ 的独立指数随机变量的和.

命题 10.2(Poisson 过程的随机选择)　设 N_t 为强度为 λ 的 Poisson 过程，如果以独立的概率 p_i 将每个跳跃时刻归入第 i 类，$i\leqslant m$，$\sum_{i=1}^{m}p_i=1$，记 $N_t^{(i)}$ 为 t 前到达的第 i 类跳跃时刻的个数. 那么$\{N_t^{(i)}: t\geqslant 0\}$是强度为 $p_i\lambda$ 的相互独立的 Poisson 过程.

命题 10.3　若 N_t 为 Poisson 过程，则在 $N_t=n$ 的条件下，各个跳跃时刻$(\tau_1,\tau_2,\cdots,\tau_n)$的联合条件分布与 n 个独立的$[0,t]$均匀随机变量的次序统计量的联合分布相同. 特别地有

(1) 在 $N_t=n$ 的条件下，τ_n 的条件分布密度为 $g_n(s)\overset{\text{def}}{=\!=}\frac{ns^{n-1}}{t^n}I_{[0,t]}(s)$，

(2) $P(\tau_n\leqslant s,N_t=n)=\dfrac{(\lambda s)^n}{n!}e^{-\lambda t}I_{[0,t]}(s)$,

(3) 在 $N_t=n$ 条件下,N_s 的条件分布与二项分布 $B\left(n,\dfrac{s}{t}\right)$ 相同.而 τ_k 的条件分布则是

$$P(\tau_k\leqslant s\mid N_t=n)=\sum_{j=k}^{n}C_n^j\left(\frac{s}{t}\right)^j\left(1-\frac{s}{t}\right)^{n-j}.$$

命题 10.4　假设$\{N_t^{(1)},t\geqslant 0\}$和$\{N_t^{(2)},t\geqslant 0\}$是独立的强度分别 λ_1,λ_2 的 Poisson 过程.对于非负整数 n 和 m,第一个过程的 n 次跳跃在第一个过程的 m 次跳跃之前发生的概率是 $\displaystyle\sum_{k=n}^{m+n-1}C_{m+n-1}^k\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{m+n-1-k}$.

10.1.2　非时齐的 Poisson 过程

定义 10.3　独立增量的计数过程 N_t,如果存在可积正函数 $\lambda(t)$ 使 $N_t-N_s\sim\exp\int_s^t\lambda(u)du$,那么 N_t 就称为**强度函数为 $\lambda(t)$ 的非时齐 Poisson 过程**.而 $\Lambda(t)=\int_0^t\lambda(s)ds$ 则称为其**均值函数**.

命题 10.5(时齐 Poisson 过程的非时齐随机选择)　设 N_t 为强度函数为 λ 的 Poisson 过程,如果在任意给定的时刻 s 以独立的概率 $p_i(s)$ 将每个跳跃时刻归入第 i 类,$i\leqslant m$,$\sum_{i=1}^m p_i(s)=1$,记 $N_t^{(i)}$ 为 t 前到达的第 i 类跳跃时刻的个数.那么$\{N_t^{(i)}:t\geqslant 0\}$是强度函数为 $\lambda\int_0^t p_i(s)ds$ 的相互独立的非时齐 Poisson 过程.

非时齐 Poisson 过程的性质　非时齐的 Poisson 过程 N_t 与 Poisson 过程的不同之处仅仅在于其强度函数不是常数 λ,而是一个时间的函数 $\lambda(t)$.对于 $s<t$,随机增量 N_t-N_s 服从参数为 $\int_s^t\lambda(u)du$ 的 Poisson 分布.由此可以通过 Poisson 分布,对非时齐的 Poisson 过程作随机模拟.

设 N_t 是一个强度为 $\lambda(t)$ 的非时齐的 Poisson 过程,它是一种记录"事故"的计数过程,将这个过程的各个计数(事故)发生的随机时刻记为 $0=\tau_0<\tau_1<\cdots<\tau_n<\cdots$.它们间仍由等式

$$\{N_t\geqslant n\}=\{\tau_n\leqslant t\}$$

相联系.与时齐的 Poisson 过程相比,这里的$\{\tau_n\}$不再独立同分布.仿照 Poisson 过程的情形可以证明如下的结论.

(1) 前 n 个发生时刻$(\tau_1,\tau_2,\cdots,\tau_n)$的联合分布密度 $g_{\tau_1,\tau_2,\cdots,\tau_n}(s_1,s_2,\cdots,s_n)$ 的表达式为

$$g_{\tau_1,\tau_2,\cdots,\tau_n}(s_1,s_2,\cdots,s_n)=\lambda(s_1)\cdots\lambda(s_n)e^{-\int_0^{s_n}\lambda(u)du}I_{\{0\leqslant s_1<\cdots<s_n\}}.$$

(2) 在已知$(\tau_1,\tau_2,\cdots,\tau_n)=(s_1,s_2,\cdots,s_n)$的条件下,随机变量$\tau_{n+1}$的条件分布密度为

$$g_{\tau_{n+1}|\tau_1,\tau_2,\cdots,\tau_n}(s_{n+1} \mid s_1,s_2,\cdots,s_n) = \lambda(s_{n+1})\mathrm{e}^{\int_{s_n}^{s_{n+1}}\lambda(u)\mathrm{d}u}I_{s_n<s_{n+1}},$$

所以,发生时刻列$\{\tau_n\}$是状态连续的非时齐的 Markov 链. 而在已知$(\tau_1,\tau_2,\cdots,\tau_n)=(s_1,s_2,\cdots,s_n)$的条件下,第$n+1$个随机间隔$T_{n+1}=\tau_{n+1}-\tau_n$的条件分布密度为

$$g_{T_{n+1}|\tau_1,\tau_2,\cdots,\tau_n}(t \mid s_1,s_2,\cdots,s_n) = \lambda(s_n+t)\mathrm{e}^{\int_{s_n}^{s_n+t}\lambda(u)\mathrm{d}u}I_{[0,\infty)}(t).$$

(3) 时刻t的计数N_t与发生时刻的联合分布是(注意这是混合型的随机向量)$P(N_t=n,\tau_1\leqslant s_1,\cdots,\tau_n\leqslant s_n)$关于$(s_1,s_2,\cdots,s_n)$的密度(称为**非时齐 Poisson 过程的样本分布**)为

$$p_{N_t,\tau_1,\tau_2,\cdots,\tau_{N_t}}(n,s_1,s_2,\cdots,s_n) = [\lambda(s_1)\cdots\lambda(s_n)I_{\{0\leqslant s_1<\cdots<s_n,n>0\}}+I_{\{n=0\}}]\mathrm{e}^{-\int_0^t\lambda(u)\mathrm{d}u}.$$

(4) 在$N_t=n$的条件下,发生时刻$(\tau_1,\tau_2,\cdots,\tau_n)$的条件分布与分布为

$$\frac{\lambda(t)}{\int_0^t\lambda(u)\mathrm{d}u}I_{[0,\infty)}(t)$$

的随机变量的次序统计量同分布(通过它可以对给定强度函数的非时齐的 Poisson 过程作计算机模拟).

(5) 设N_t是以$\lambda(t)$为强度函数的非时齐的 Poisson 过程,那么$\widetilde{N}_t \stackrel{\text{def}}{=\!=} N_t-\int_0^t\lambda(s)\mathrm{d}s$是鞅. 因此,随机过程$\Lambda_t\stackrel{\text{def}}{=\!=}\int_0^t\lambda(s)\mathrm{d}s$也称为非时齐的 Poisson 过程$N_t$的补偿函数.

在一些参数模型的建模问题中,常常用带有未知参数的强度函数$\lambda(t)$的非时齐的 Poisson 过程建模. 这时需要用实测数据来估计这些未知参数,例如用最大似然估计,于是需将观测到的样本置入分布中去,为此,我们需要**非时齐的 Poisson 过程的似然函数.**

对于非时齐的 Poisson 过程,我们有(由过程的一个样本给出的)似然函数: 在$N_t=0$时

$$p(N_t,\tau_1,\cdots,\tau_{N_t}) = \mathrm{e}^{-\int_0^t\lambda(u)\mathrm{d}u},$$

而在$N_t>0$时

$$p(N_t,\tau_1,\cdots,\tau_{N_t}) = \mathrm{e}^{-\int_0^t\lambda(u)\mathrm{d}u+\sum_{k=1}^{N_t}\ln\lambda(\tau_k)} = \mathrm{e}^{-\int_0^t\lambda(u)\mathrm{d}u+\int_0^t\ln\lambda(u)\mathrm{d}N_u}.$$

下面是典型的建模的例子.

例 10.1 (1) 放射源发射的光子数目可以用强度为

$$\lambda(t,\alpha,\beta) = \alpha\mathrm{e}^{-\beta t} \quad (\alpha,\beta>0)$$

的非时齐的 Poisson 过程来建模. 它模拟了放射强度的衰变情况.

(2) 在核医疗中使用的光脉冲序列通常用强度为

$$\lambda(t,\gamma,\alpha_1,\alpha_2,\cdots,\alpha_n,\beta_1,\beta_2,\cdots,\beta_n) = \gamma+\sum_{k=1}^n\alpha_i\mathrm{e}^{-\beta_i t} \quad (\alpha_k,\beta_k,\gamma>0)$$

的非时齐的 Poisson 过程来模拟并建模.

(3) 调幅、调相与调频为 f_m 的脉冲光源则可以分别用强度为

$$\lambda(t,\gamma,\alpha,\beta)=\gamma+\alpha\beta\mid S(t)\mid^2,$$
$$\lambda(t,\gamma,\alpha,\beta)=\gamma+\alpha\mid S(t-\beta)\mid^2$$

与

$$\lambda(t,\gamma,\alpha,\beta,m)=\gamma+\alpha(1+m\cos[2\pi(f_m+\beta)t])\quad(\mid m\mid<1)$$

的非时齐的 Poisson 过程模拟.

注　在用例 10.1 模型建模时,需要用过程的一段观测轨道估计未知参数.而将似然函数取得最大值的待定参数的值作为估计,就是最大似然估计.

10.1.3　复合 Poisson 过程与非时齐的复合 Poisson 过程

定义 10.4　设 $\{\eta_n\}$ 独立同分布序列,而 N_t 是一个与它独立的强度为 λ 的 Poisson 过程,则随机过程 $\zeta_t=\eta_1+\eta_2+\cdots+\eta_{N_t}$ 称为强度为 λ 的**复合 Poisson 过程**.如果 N_t 是一个与它独立的强度函数为 $\lambda(t)$ 的非时齐 Poisson 过程,则随机过程 $\zeta_t=\eta_1+\eta_2+\cdots+\eta_{N_t}$ 称为强度函数为 $\lambda(t)$ 的**非时齐的复合 Poisson 过程**.

复合 Poisson 过程 ζ_t 的特征函数为

$$E\mathrm{e}^{\mathrm{i}\alpha\zeta_t}=\mathrm{e}^{\lambda(E\mathrm{e}^{\mathrm{i}\alpha\eta_1}-1)t}.$$

10.2　对非时齐的 Poisson 过程的随机积分与对 Poisson 点过程的随机积分

10.2.1　对非时齐的 Poisson 过程 N_t 的随机积分

利用非时齐的 Poisson 过程在长度为 h 的区间上的增量大于 1 的概率为 $o(h)$,如下的引理是显然的.

引理 10.6　对于非时齐的 Poisson 过程 N_t 及 $(0,t]$ 的一个划分:

$$0=t_0^{(n)}<\cdots<t_j^{(n)}<t_{j+1}^{(n)}<\cdots<t_n^{(n)}=t,\quad \max_j\{t_{j+1}^{(n)}-t_j^{(n)}\}\to 0(n\to+\infty),$$

有(由于非时齐的 Poisson 过程或者不跳跃,或者跳跃为 1)

$$\lim_{n\to+\infty}\sum_{k=1}^{n}(N_{t_k^{(n)}}-N_{t_{k-1}^{(n)}})^2=N_t. \tag{10.1}$$

即

$$(\mathrm{d}N_t)^2=\mathrm{d}N_t\qquad(\text{只取 0 或 1}). \tag{10.1$'$}$$

与 Brown 运动相比,这里是 $(\mathrm{d}N_t)^2=\mathrm{d}N_t$,而在 Brown 运动是 $(\mathrm{d}B_t)^2=\mathrm{d}t$.

定义 10.5　假设 N_t 是一个强度函数为 $\lambda(t)$ 的非时齐 Poisson 过程,相继的事件到达的时间为 $\{\tau_n\}$.对于 (N_t) 可知的,且轨道具有左极限的随机过程 Ψ_t,对于固定的样本 ω,

定义对非时齐 Poisson 过程的按样本的随机积分为

$$\int_0^t \Psi_{s-} \mathrm{d}N_s = \lim_{n\to+\infty} \sum_{k=1}^{n} \Psi_{t_{k-1}^{(n)}} (N_{t_k^{(n)}} - N_{t_{k-1}^{(n)}}) = \begin{cases} \sum_{n=1}^{N_t} \Psi_{\tau_n}, & N_t \geqslant 1, \\ 0, & N_t = 0, \end{cases} \tag{10.2}$$

如果右方的量概率为 1 地收敛. 注意这里被积函数用的是左极限,原因在于积分和中被积函数取的是区间左端点的值.

这时不定限积分$\int_0^t \Psi_{s-} \mathrm{d}N_s$ 也是(N_t)可知的随机过程.

另一方面,对鞅$\widetilde{N}_t$ 的随机积分是 Ito 积分的非时齐的 Poisson 版本,而且有

$$\int_0^t \Psi_s \mathrm{d}N_s = \int_0^t \Psi_s \mathrm{d}\,\widetilde{N}_s + \int_0^t \Psi_s \lambda(s)\mathrm{d}s.$$

关于时齐的 Poisson 过程的随机积分有许多与 Ito 积分相仿的性质. 如:

(1) $E\left(\int_0^t f(s)\mathrm{d}N_s\right) = \int_0^t f(s)\lambda(s)\mathrm{d}s$,

(2) $\mathrm{var}\left(\int_0^t f(s)\mathrm{d}N_s\right) = \int_0^t f(s)^2\lambda(s)\mathrm{d}s$.

证明 利用非时齐 Poisson 过程的结论(4)以及对称性,可以得到

$$\sum_{k=1}^{n} E[f(\tau_k) \mid N_t = n] = E\left(\sum_{k=1}^{n} f(\eta_k^*)\right) = E\left(\sum_{k=1}^{n} f(\eta_k)\right) = \sum_{k=1}^{n} Ef(\eta_k)$$
$$= n\int_0^t f(u)\frac{\lambda(u)}{\int_0^t \lambda(s)\mathrm{d}s}\mathrm{d}u.$$

于是

$$E\left(\int_0^t f(s)\mathrm{d}N_s\right) = \sum_{n=0}^{\infty}\sum_{k=1}^{n} E[f(\tau_k) \mid N_t = n]P(N_t = n)$$
$$= EN_t\int_0^t f(u)\frac{\lambda(u)}{\int_0^t \lambda(s)\mathrm{d}s}\mathrm{d}u = \int_0^t f(u)\lambda(u)\mathrm{d}u.$$

此即第一个等式. 另一个的证明是类似的,只是稍复杂一些.

随机函数对 Poisson 过程的积分

例 10.2

$$\int_0^t N_{s-}\mathrm{d}N_s = \frac{1}{2}N_t^2 - \frac{1}{2}N_t.$$

证明 仿照 Brown 运动情形的计算,可由(10.2)式直接算得.

关于非时齐 Poisson 过程的随机积分与非时齐 Poisson 过程的关系,类似于 Ito 积分与

Brown 运动的关系，所以这种积分除了可加性等普适性质外，(1)与(2)还可以加强为：

(1) $\int_0^t \Psi_{s-}\mathrm{d}N_s - \int_0^t \Psi_s\lambda(s)\mathrm{d}s$ 是鞅(类似于不定限 Ito 积分是鞅).

(2) $\left(\int_0^t \Psi_{s-}\mathrm{d}N_s\right)^2 - \int_0^t \Psi_s^2\lambda(s)\mathrm{d}s$ 也是鞅(类似于 Ito 积分的平方可积鞅).

10.2.2　与非时齐的复合 Poisson 过程相系的 Poisson 点过程(用时间、空间积分表示)

(1) 将非时齐复合 Poisson 过程表为非时齐 Poisson 过程的积分(用时间积分表示)

利用对于非时齐的 Poisson 过程的随机积分的定义可得下述结论.

设 N_t 为非时齐的 Poisson 过程，$\{\eta_n\}$ 为与之独立的独立同分布随机变量序列. 而 $\xi_t = \sum_{k=1}^{N_t}\eta_k$ 是非时齐的复合 Poisson 过程，则它有以下的积分表示式

$$\xi_t = \int_0^t \tilde{\eta}_s \mathrm{d}N_s, \tag{10.3}$$

其中 $\tilde{\eta}_s = \begin{cases}\eta_n, s=\tau_n, \\ 0, 其他.\end{cases}$

(2) 将非时齐复合 Poisson 过程表为 Poisson 点过程的积分(用空间积分表示)

先讨论简单的情形. 假定 $\eta_n \sim \begin{pmatrix} v_1 & \cdots & v_k & \cdots \\ p_1 & \cdots & p_k & \cdots \end{pmatrix}$，非时齐的复合 Poisson 过程 $\xi_t = \sum_{k=1}^{N_t}\eta_k$. 将过程在时刻 t 以前取值 v 的累计次数记为

$$N_t(\{v\}) = \{对\ s \leqslant t, \eta_s\ 取值\ v\ 的次数\}. \tag{10.4}$$

对于固定的值 v，$N_t(\{v\})$可以看成对于非时齐的 Poisson 过程 N_t 的随机选择. 由随机选择定理可知，$N_t(\{v_k\})$是强度函数为 $p_k\lambda(t)$的非时齐 Poisson 过程.

由非时齐的复合 Poisson 过程的定义，直观地看，对于 $v = v_1, v_2, \cdots, v_k, \cdots$，$\{N_t(\{v\})\}$是一系列相互独立的非时齐的 Poisson 过程. 而且有

$$\xi_t = \sum_k v_k N_t(\{v_k\}).$$

因为 Poisson 过程序列$\{N_t(\{v\})\}(v=v_1, v_2, \cdots, v_k, \cdots)$与空间的取值点列$\{v_1, v_2, \cdots, v_k, \cdots\}$相系，所以称$\{N_t(\{v\})\}$为 **Poisson 点过程**.

Poisson 点过程的概念比 Poisson 过程复杂，它包含时空两个参数(t, v)，而且对于不同的 v，作为时间 t 的随机过程是相互独立的 Poisson 过程.

表达式(10.4)的直观含义是，如果将非时齐的复合 Poisson 过程 ξ_t 看成在时刻 t 的总的随机积累，那么它是时刻 t 前取大小不同的固定值的随机积累的总和.

这个思想可以拓展到一般的非时齐的复合 Poisson 过程 ξ_t，即随机变量 η_k 不必局限于取离散值，而是可以取任意值(甚至可取负值)的情形. 这时 ξ_t 也可以取任意的值. 它在时刻 t 前取值于区间 $(a,b]$ 的次数，是一个随机过程，记之为 $N_t((a,b])$. 因为 ξ_t 当且仅当在此非时齐的 Poisson 过程 N_t 的事件列 τ_n 上跳跃，所以这个次数是一个有限的(但是是随机的数)，它们满足

(P.1) 由随机选择定理，$N_t((a,b])$ 是强度为 $P(\eta_1\in(a,b])\lambda(t)$ 的非时齐的 Poisson 过程.

(P.2) 由随机选择定理，若区间 $(a_i,b_i]$ $(i=1,2,\cdots,m)$ 两两不交，则随机过程 $N_t((a_i,b_i])$ $(i=1,2,\cdots,m)$ 是相互独立的.

(P.3) 对于 $a<b<c$ 有可加性：$N_t((a,b])+N_t((b,c])=N_t((a,c])$.

再记

$$N_t(v)\overset{\text{def}}{=\!=}N_t((-\infty,v]). \tag{10.5}$$

我们将它叙述为如下的略广一些(不仅仅限于非时齐的复合 Poisson 过程)的定义.

警示 这里的记号 $N_t(v)$ 与前面定义的记号 $N_t(\{v\})$ 的含义是不同的.

定义 10.6 依赖于实数值 v 的计数随机过程族 $\{N_t(v): t\geqslant 0, -\infty<v<+\infty\}$，称为 **Poisson 点过程**，如果对于 $N_t((a,b])\overset{\text{def}}{=\!=}N_t(b)-N_t(a)$ 满足以上的条件(P.2)，(P.3)，以及如下的(P.1)$'$：

(P.1)$'$ 存在单调递增函数 $F(v)$，使 $F(-\infty)=0$ 而且 $N_t((a,b])$ 是强度函数为 $[F(b)-F(a)]\lambda(t)$ 的非时齐的 Poisson 过程.

此时 $F(v)\lambda(t)$ 称为 **Poisson 点过程的补偿函数**.

例 10.3 由非时齐的复合 Poisson 过程 ξ_t 相系的随机过程族 $\{N_t(v): t\geqslant 0, -\infty<v<+\infty\}$ 是 Poisson 点过程，其中 $N_t(v)$ 表示此非时齐的复合 Poisson 过程在时刻 t 前取值于区间 $(-\infty,v]$ 的次数. 易见此 Poisson 点过程的补偿函数是 $F_\eta(v)\lambda(t)$，其中 F_η 是 η_n 的分布函数，即 $\widetilde{N}_t(v)=N_t(v)-F_\eta(v)\lambda(t)$ 是 (N_t) 鞅.

反之，若 Poisson 点过程的补偿函数 $F(v)\lambda(t)$ 满足 $F(+\infty)-F(-\infty)=C<+\infty$，则存在以 $C\lambda(t)$ 为强度函数的非时齐的 Poisson 过程，及对应于赋值随机变量 η_n 的分布函数为 $F_{\eta_n}(v)\overset{\text{def}}{=\!=}\dfrac{1}{C}F(v)$ 的非时齐的复合 Poisson 过程，使此 Poisson 点过程由此非时齐的复合 Poisson 过程生成.

特别地，对于非时齐的 Poisson 过程 N_t 有

$$N_t(\{k\})=I_{\{\tau_k\leqslant t<\tau_{k+1}\}},\quad N_t(v)=\sum_{k\geqslant 1}N_t(\{k\})I_{\{k\leqslant v<k+1\}}.$$

直观地，非时齐的复合 Poisson 过程 ξ_t 可以看成在时刻 t 的大小(可以取负值)不同的随机积累的总和，即对于划分

$$\cdots < v_i^{(n)} < v_{i+1}^{(n)} < \cdots < v_{m_n}^{(n)}, \quad v_0^{(n)} = -n, \quad v_{m_n}^{(n)} = n,$$
$$\max_i \{v_{i+1}^{(n)} - v_i^{(n)}\} \to 0 \quad (n \to +\infty), \quad v_i^{(n)} < v_i^{(n)\prime} \leqslant v_{i+1}^{(n)},$$

有

$$\xi_t \approx \sum_i v_i^{(n)\prime} N_t((v_i^{(n)}, v_{i+1}^{(n)}]) = \sum_i v_i^{(n)\prime} [N_t(v_{i+1}^{(n)}) - N_t(v_i^{(n)})].$$

当 $n \to +\infty$ 时，可以证明上式右方依概率收敛，而且不依赖于划分及 $v_i^{(n)\prime}$ 的取法. 将此依概率收敛的极限随机变量记为 $\int v N_t(\mathrm{d}v)$. 于是就可以用 Poisson 点过程表达非时齐复合 Poisson 过程 ξ_t 为

$$\xi_t = \int v N_t(\mathrm{d}v). \tag{10.3$'$}$$

(10.3)′式比(10.3)式更直观.

10.2.3 将非时齐复合 Poisson 过程表为时空 Poisson 点过程的积分(用时间、空间积分表示)

还可以从另一种角度考虑与非时齐的复合 Poisson 过程 ξ_t 相系的 Poisson 点过程. 改记

$$\mu(t, v) = N_t(v), \tag{10.6}$$

并且将它视为时空双参数的计数随机变量族，简称为**时空 Poisson 点过程**. 对于 $0 \leqslant s < t$，$a < b$，定义

$$\mu((s,t] \times (a,b]) \overset{\text{def}}{=} \mu(t,b) - \mu(t,a) - \mu(s,b) + \mu(s,a). \tag{10.7}$$

容易验证：

(S.1) Poisson 性：$\mu((s,t] \times (a,b]) \sim \text{Poisson}_{(\lambda(t)-\lambda(s))(F_{\eta_1}(b) - F_{\eta_1}(a))}$，

其中 F_{η_1} 是这个非时齐复合 Poisson 过程相系的赋值随机变量 η_k 们的公共分布函数.

(S.2) 独立性：若 $A_i = (s_i, t_i] \times (a_i, b_i]$ $(i = 1, 2, \cdots, m)$，且 A_i 们两两不相交，则 $\mu(A_1), \mu(A_2), \cdots, \mu(A_m)$ 彼此独立.

(S.3) 可加性：若 $A_i = (s_i, t_i] \times (a_i, b_i]$ $(i = 1, 2, 3)$，且 $A_2 \cap A_3 = 0$，$A_1 = A_2 \cup A_3$，则

$$\mu(A_1) = \mu(A_2) + \mu(A_3).$$

我们给出一个一般的(不仅仅限于非时齐的复合 Poisson 过程)定义.

定义 10.7 计数随机过程族 $\{\mu(t,v): t \geqslant 0, -\infty < v < +\infty\}$，称为**时空 Poisson 点过程**，如果条件(S.1)′，(S.2)，(S.3)满足，其中 $\mu((s,t] \times (a,b])$ 由(10.7)式定义，这里

(S.1)′存在 F 单调递增，$F(-\infty) = 0$，使

$$\mu((s,t] \times (a,b]) \sim \text{Poisson}_{(\lambda(t)-\lambda(s))(F(b)-F(a))}.$$

$F(v)\lambda(t)$ 称为时空 Poisson 点过程的**补偿函数**.

注意 $\mu(\mathrm{d}t, \mathrm{d}v)$ 是计数过程，所以显然有(只能取 0 或 1)

$$(\mu(\mathrm{d}t, \mathrm{d}v))^2 = \mu(\mathrm{d}t, \mathrm{d}v).$$

注 1　这里只要求 F 递增，并未假定 $F(\infty)<+\infty$. 我们看到只有在 $F(+\infty)=+\infty$ 的时空 Poisson 点过程才不是由非时齐的 Poisson 过程所生成的.

注 2　一般地，空间变量 v 可以是多维的. 只需对定义作一些必要的修改.

例 10.4　由非时齐的复合 Poisson 过程所定义的随机过程族

$$\{\mu(t,v):\ t\geqslant 0,\ -\infty<v<+\infty\}$$

是时空 Poisson 点过程，其中 $\mu(t,v)=N_t(v)$ 表示此非时齐的复合 Poisson 过程在时刻 t 前的跳跃值在区间 $(-\infty,v]$ 中的次数.

类似地，若时空 Poisson 点过程的补偿函数 $F(v)\lambda(t)$ 满足 $F(+\infty)-F(-\infty)=C<+\infty$，则存在以 $C\lambda(t)$ 为强度函数的非时齐的 Poisson 过程，及对应于赋值随机变量 η_n 的分布函数为 $F_{\eta_n}(v)=\dfrac{1}{C}F(v)$ 的非时齐的复合 Poisson 过程，使这个时空 Poisson 点过程是由此非时齐的复合 Poisson 过程所生成.

对于非时齐的 Poisson 过程 N_t，有

$$\mu((s,t]\times(a,b])=\sum_{a<k\leqslant b}I_{\{s<\tau_k\leqslant t<\tau_{k+1}\}},$$

$$N_t=\mu((-\infty,t]\times(0,+\infty))=\sum_k I_{\{\tau_k\leqslant t\}}=\sup\{k:\ \tau_k\leqslant t\}.$$

对于划分

$$\cdots<v_i^{(n)}<v_{i+1}^{(n)}<\cdots<v_{m_n}^{(n)},\quad v_0^{(n)}=-n,\quad v_{m_n}^{(n)}=n,$$

$$0=s_0^{(n)}<\cdots<s_j^{(n)}<s_{j+1}^{(n)}<\cdots<s_n^{(n)}=t,$$

$$\max_{i,j}\{v_{i+1}^{(n)}-v_i^{(n)},s_{j+1}^{(n)}-s_j^{(n)}\}\to 0\quad(n\to+\infty),\quad v_i^{(n)}<v_i^{(n)\prime}\leqslant v_{i+1}^{(n)}.$$

有

$$\xi_t\approx\sum_i v_i^{(n)\prime}[\mu(t,v_{i+1}^{(n)})-\mu(t,v_i^{(n)})]=\sum_{i,j}v_i^{(n)}\mu((s_i^{(n)},s_{i+1}^{(n)}]\times(v_j^{(n)},v_{j+1}^{(n)}]).$$

当 $n\to+\infty$ 时，可以证明上式右方依概率收敛. 将此依概率收敛的极限随机变量记为 $\int_0^t\int_{-\infty}^{+\infty}v\mu(\mathrm{d}s,\mathrm{d}v)$. 于是就可以用时空 Poisson 点过程表达非时齐复合 Poisson 过程 ξ_t

$$\xi_t=\int_0^t\int_{-\infty}^{+\infty}v\mu(\mathrm{d}s,\mathrm{d}v).\tag{10.3$''$}$$

事实上，(10.3)′式是(10.3)″式的边际积分形式.

10.3　对时空 Poisson 点过程 $\mu(t,v)$ 的随机积分

10.3.1　对时空 Poisson 点过程 $\mu(t,v)$ 的随机积分的必要性

如果用对 Poisson 过程的积分代替 Ito 积分，那么 Ito 过程对应的就是以下的 Ito-Skorohod 过程.

定义 10.8(Ito-Skorohod 过程)　形如

$$\xi_t = \int_0^t \Phi_s \mathrm{d}s + \int_0^t \Psi_{s-} \mathrm{d}N_s \quad (\text{简记为 } \mathrm{d}\xi_t = \Phi_t \mathrm{d}t + \Psi_{t-} \mathrm{d}N_t) \tag{10.8}$$

的随机过程 ξ_t,称为 **Ito-Skorohod 过程**,其中 N_t 是非时齐的 Poisson 过程.

然而,Ito-Skorohod 过程的连续可微函数的复合,远比 Ito 过程的复合函数的 Ito 公式复杂,它们将不再是 Ito-Skorohod 过程.我们直观地分析如下:

如果 $F(x)$ 连续可微,ξ_t 是 Ito-Skorohod 过程 $\mathrm{d}\xi_t = \Phi_t \mathrm{d}t + \Psi_{t-} \mathrm{d}N_t$,则 $\mathrm{d}F(\xi_t)$ 只在 t 是非时齐的 Poisson 过程的跳跃时刻时可能有非零值.于是

$$\begin{aligned}
\mathrm{d}F(\xi_{t-}) &= F(\xi_{t-} + \mathrm{d}\xi_t) - F(\xi_{t-}) = F(\xi_{t-} + \Phi_t \mathrm{d}t + \Psi_{t-} \mathrm{d}N_t) - F(\xi_{t-}) \\
&= \{F(\xi_{t-} + \Phi_t \mathrm{d}t + \Psi_{t-} \mathrm{d}N_t) - F(\xi_{t-} + \Psi_{t-} \mathrm{d}N_t)\} + \{F(\xi_{t-} + \Psi_{t-} \mathrm{d}N_t) - F(\xi_{t-})\} \\
&\approx F'(\xi_{t-})\Phi_t \mathrm{d}t + \{F(\xi_t + \Psi_{t-} \mathrm{d}N_t) - F(\xi_t)\} \\
&\approx F'(\xi_{t-})\Phi_t \mathrm{d}t + \sum_k [F(\xi_{t-} + \Psi_{t-}) - F(\xi_{t-})] I_{\{\tau_k = t\}} \mathrm{d}N_t,
\end{aligned}$$

其中 $I_{\{\tau_k=t\}}$ 的出现使得用 Ito-Skorohod 过程表达 $\mathrm{d}F(\xi_t)$ 不够清晰(清晰的表达应该只出现函数 F 与被积过程 Φ_t, Ψ_{t-} 和 $\mathrm{d}N_t$,不应该出现 $I_{\{\tau_k=t\}}$),N_t 是一个强度函数为 $\lambda(t)$ 的非时齐 Poisson 过程,相继的事件到达的时间为 $\{\tau_n\}$.要想使右边的和有清晰的解释,可以借助对时空 Poisson 点过程的积分(凡是涉及 $\tau_k \leqslant t$ 的和项,都适于用对时空 Poisson 点过程的积分).表达为

$$\mathrm{d}F(\xi_t) = F'(\xi_t)\Phi_t \mathrm{d}t + \int [F(\xi_{t-} + \Psi_{t-}) - F(\xi_{t-})]\mu(\mathrm{d}t, \mathrm{d}v).$$

于是我们看到为了得到 Ito-Skorohod 过程的对光滑的函数复合的良好表达,应该使用对时空的 Poisson 点过程的随机积分.

10.3.2　对时空 Poisson 点过程 $\boldsymbol{\mu}(t, v)$ 的随机积分的含义

对于带有参数 v 的 $(\mu(t,v);\ \forall v)$ 可知的,且对 t 有左极限的随机过程 $\Psi_t(v)$ 及划分:

$$\cdots < v_i^{(n)} < v_{i+1}^{(n)} < \cdots < v_{m_n}^{(n)}, \quad v_0^{(n)} = -n, \quad v_{m_n}^{(n)} = n,$$

$$0 = s_0^{(n)} < \cdots < s_j^{(n)} < s_{j+1}^{(n)} < \cdots < s_n^{(n)} = t,$$

$$\max_{i,j}\{v_{i+1}^{(n)} - v_i^{(n)}, s_{j+1}^{(n)} - s_j^{(n)}\} \to 0 \quad (n \to +\infty), \quad v_i^{(n)} < v_i^{(n)\prime} \leqslant v_{i+1}^{(n)}.$$

定义对时空 Poisson 点过程的随机积分为(在概率收敛的含义下)

$$\int_0^t \int \Psi_{s-}(v)\mu(\mathrm{d}s, \mathrm{d}v) = \lim_{n \to +\infty} \sum_{i,j} \Psi_{s_{j-1}^{(n)}}(v_i^{(n)})\mu((s_i^{(n)}, s_{i+1}^{(n)}] \times (v_j^{(n)}, v_{j+1}^{(n)}]). \tag{10.9}$$

对时空 Poisson 点过程的随机积分有与 Ito 积分类似的许多性质,如可加性,以及

(1) $\int_0^t \int \Psi_{s-}(v)\mu(\mathrm{d}s, \mathrm{d}v) \int_0^t - \int \Psi_s(v)\lambda(s)\mathrm{d}s\mathrm{d}F(v)$ 是鞅.

(2) $\left[\int_0^t \int \Psi_{s-}(v)\mu(\mathrm{d}s, \mathrm{d}v)\right]^2 - \int_0^t \int \Psi_s^2(v)\lambda(s)\mathrm{d}s\mathrm{d}F(v)$ 也是鞅.

设对应于 $N_t \stackrel{\mathrm{def}}{=} \int \mu(t, \mathrm{d}v)$ 的事件发生时刻列为 $\{\tau_n\}$. 在 F 是分布函数时，这个时空 Poisson 点过程由某个非时齐的复合 Poisson 过程 ξ_t 生成. 设这个非时齐的复合 Poisson 过程在 τ_n 的赋值为 η_n，正好有

$$\int_0^t \int \Psi_{s-}(v)\mu(\mathrm{d}s, \mathrm{d}v) = I_{\{N_t \geqslant 1\}} \sum_{n=1}^{N_t} \Psi_{\tau_n}(\eta_n).$$

10.4 以 Poisson 过程或时空 Poisson 点过程驱动的随机微分方程与 Poisson 随机微积分的复合函数的 Ito 公式

10.4.1 以时空 Poisson 过程驱动的随机微分方程

时空 Poisson 过程驱动的随机积分方程

$$\xi_t = \xi_0 + \int_0^t b(s, \xi_s)\mathrm{d}s + \int_0^t \int g(s, \xi_{s-}, v)\mu(\mathrm{d}s, \mathrm{d}v) \tag{10.10}$$

称为**时空 Poisson 过程驱动的随机微分方程**，其中 $b(t,x)$，$g(t,x,v)$ 都是 (t,x) 的连续函数. 简记为如下的微分形式

$$\mathrm{d}\xi_t = b(t, \xi_t)\mathrm{d}t + \int g(t, \xi_{t-}, v)\mu(\mathrm{d}t, \mathrm{d}v). \tag{10.10$'$}$$

而由非时齐的 Poisson 过程驱动的随机微分方程

$$\xi_t = \xi_0 + \int_0^t b(s, \xi_s)\mathrm{d}s + \int_0^t g(s, \xi_{s-}, v)\mathrm{d}N_s \tag{10.11}$$

是方程(10.10)的特例.

如果一个随机过程 ξ_t 满足方程(10.10)，就称为方程(10.10)的解. 这时的增量

$$\xi_t - \xi_{t-} = \int g(t, \xi_{t-}, v)\mu(\{t\}, \mathrm{d}v)$$

未必等于 0. 所以方程(10.10)的解的轨道与 Ito 方程的解的轨道很不同，Ito 方程的解是连续的，而是方程(10.10)的解是可以有跳跃的.

此外，与 Ito 方程完全类似地，可以证明时空 Poisson 过程驱动的随机微分方程的解也是 Markov 过程. 于是时空 Poisson 过程驱动的随机微分方程就成为一类轨道可以有间断的 Markov 过程的建模的有力工具. 例如，可以用它改进风险证券的 Black-Scholes 模型，使之可以包容市场的突然变化. 例如用时空 Poisson 过程和与其独立的 Brown 运动共同驱动的随机微分方程

$$\mathrm{d}\xi_t = b(t, \xi_t)\mathrm{d}t + \sigma(t, \xi_t)\mathrm{d}B_t + \int g(t, \xi_{t-}, v)\mu(\mathrm{d}t, \mathrm{d}v) \tag{10.12}$$

建模.

例 10.5(乘积的 Ito 公式与分部积分)　对于随机微分方程

$$\mathrm{d}\boldsymbol{\xi}_t=\int \boldsymbol{g}(t,\boldsymbol{\xi}_{t-},\boldsymbol{v})\mu(\mathrm{d}t,\mathrm{d}\boldsymbol{v})$$

及 $\boldsymbol{F}(t,\boldsymbol{x})$有

$$\mathrm{d}\boldsymbol{F}(t,\boldsymbol{\xi}_t)=\int[\boldsymbol{F}(t,\boldsymbol{\xi}_{t-}+\boldsymbol{g}(t,\boldsymbol{\xi}_{t-},\boldsymbol{v}))-\boldsymbol{F}(t,\boldsymbol{\xi}_{t-})]\mu(\mathrm{d}t,\mathrm{d}v).$$

特别地,对于

$$\mathrm{d}\xi_t^{(i)}=\int g_i(t,\xi_{t-}^{(i)},\boldsymbol{v})\mu(\mathrm{d}t,\mathrm{d}\boldsymbol{v})\quad(i=1,2)$$

有乘积公式

$$\begin{aligned}\mathrm{d}(\xi_t^{(1)}\xi_t^{(2)})=\int[&\xi_t^{(1)}g_2(t,\xi_{t-}^{(2)},\boldsymbol{v})+\xi_t^{(2)}g_1(t,\xi_{t-}^{(1)},\boldsymbol{v})\\&+g_1(t,\xi_{t-}^{(1)},\boldsymbol{v})g_2(t,\xi_{t-}^{(2)},\boldsymbol{v})]\mu(\mathrm{d}t,\mathrm{d}\boldsymbol{v}).\end{aligned}$$

即分部积分

$$\mathrm{d}(\xi_t^{(1)}\xi_t^{(2)})=\xi_t^{(1)}\mathrm{d}\xi_t^{(2)}+\xi_t^{(2)}\mathrm{d}\xi_t^{(1)}+\int g_1(t,\xi_{t-}^{(1)},\boldsymbol{v})g_2(t,\xi_{t-}^{(2)},\boldsymbol{v})\mu(\mathrm{d}t,\mathrm{d}\boldsymbol{v}).\tag{10.13}$$

(10.13)式可以推广,即对于

$$\mathrm{d}\boldsymbol{\xi}_t=\boldsymbol{b}(t,\boldsymbol{\xi}_t)\mathrm{d}t+\int \boldsymbol{g}(t,\boldsymbol{\xi}_{t-},\boldsymbol{v})\mu(\mathrm{d}t,\mathrm{d}\boldsymbol{v})$$

的情形仍然正确. 从而得到一般的分部积分公式.

下面叙述存在唯一性定理,其证明可以在随机微分方程的书的有关带跳的随机微分方程的章节中找到,我们略去.

定理 10.7(时空 Poisson 过程驱动的随机微分方程解的存在唯一性)　若 $\lambda(t)$有界,$g(t,x,v)$连续,且在$[0,T]$上满足

$$b(t,x)^2+\sigma(t,x)^2+g(t,x,v)^2\leqslant C(1+x^2),\tag{10.14}$$

$$|b(t,x)-b(t,y)|+|\sigma(t,x)-\sigma(t,y)|+|g(t,x,v)-g(t,y,v)|^2\leqslant C|x-y|^2.\tag{10.15}$$

那么随机微分方程(10.12)存在唯一的解 ξ_t,而且它是 Markov 过程.

例 10.6　设 $f(t)$是数值函数,N_t 是非时齐的 Poisson 过程. 考虑随机微分方程

$$\mathrm{d}\xi_t=\xi_t f(t)\mathrm{d}N_t,\quad \xi_0=1.$$

它有唯一解. 可以用 Ito 公式验证这个解有表达式

$$\xi_t=\mathrm{e}^{\int_0^t\ln(1+f(s))\mathrm{d}N_s}=\prod_{n=1}^{N_t}(1+f(\tau_n)).\tag{10.16}$$

10.4.2　一般的随机微分方程的解的 Ito 公式

定理 10.8(一般的 Ito 公式)　若 $F(t,x)$对 t 连续可微,对 x 二阶连续可微,ξ_t 是方程(10.13)的解,则

$$\mathrm{d}F(t,\xi_{t-})=\left[F'_t+b(t,\xi_t)F'_x+\frac{1}{2}\sigma^2(t,\xi_t)F''_{xx}\right]\mathrm{d}t+\sigma(t,\xi_t)F'_x\mathrm{d}B_t$$
$$+\int[F(t,\xi_{t-}+g(t,\xi_{t-},v))-F(t,\xi_{t-})]\mu(\mathrm{d}t,\mathrm{d}v).$$

同样,在多维情形也有对应的定理.

例 10.7 随机微分方程

$$\mathrm{d}\zeta_t=\zeta_t(g(t)\mathrm{d}B_t+f(t)\mathrm{d}N_t),\quad \zeta_0=1$$

的唯一解是

$$\zeta_t=\mathrm{e}^{\int_0^t g(s)\mathrm{d}B_s-\frac{1}{2}\int_0^t g(s)^2\mathrm{d}s}\prod_{n=1}^{N_t}(1+f(\tau_n)). \tag{10.17}$$

10.5 常见的条件 Poisson 过程

10.5.1 自激点过程(条件强度过程,绝对概率,事件到达时间联合分布,样本分布)

(1) 自激点过程的定义

定义 10.9 如果计数过程 N_t 满足:$N_0=0$,且对于小时间 h 有

$$P(N_{t+h}-N_t=1\mid N_t=0)=\lambda(t,0)h+o(h),$$
$$P(N_{t+h}-N_t=1\mid N_t=k,\tau_1=s_1,\cdots,\tau_k=s_k)=\lambda(t,N_t;\ s_1,\cdots,s_k)h+o(h)\quad(k\geqslant 1),$$
$$P(N_{t+h}-N_t\geqslant 2\mid N_t=k,\tau_1=s_1,\cdots,\tau_k=s_k)=o(h)\quad(k\geqslant 1). \tag{10.18}$$

则称为**自激点过程**.而随机过程

$$\lambda_t\overset{\text{def}}{=\!=}\begin{cases}\lambda(t,0), & t<\tau_1,\\ \lambda(t,N_t;\ \tau_1,\cdots,\tau_{N_t}), & \tau_{N_t}\leqslant t<\tau_{N_t+1}\end{cases} \tag{10.19}$$

称为其**强度过程**.

自激点过程 N_t 称为***m*-记忆的**,如果

$$\lambda(t,N_t;\ \tau_1,\cdots,\tau_{N_t})=\lambda(t,N_t;\ \tau_{N_t-m+1},\cdots,\tau_{N_t})$$

即强度过程只依赖最近发生的 m 个事件的时刻,而与 $\tau_1,\cdots,\tau_{N_t-m}$ 无关.

最简单的是 $\lambda_t=\lambda(t,N_t)$ 的情形.这是无记忆(0-记忆)的自激点过程情形.即下面的定义。

定义 10.10 计数过程 N_t 称为**无记忆的自激点过程**,如果 N_t 满足:$N_0=0$,且存在正值函数 $\lambda(t,m)$,使对于小的 h 有

$$P(N_{t+h}-N_t=1\mid N_t,\tau_1,\cdots,\tau_{N_T})=\lambda(t,N_t)h+o(h),$$
$$P(N_{t+h}-N_t\geqslant 2\mid N_t,\tau_1,\cdots,\tau_{N_t})=o(h).$$

由此可见,无记忆的自激点过程是条件非时齐 Poisson 过程.

例 10.8 年龄 x 的 n 个人的群体在时刻 t 以前的死亡人数(死亡计数过程)N_t 可以

用无记忆的自激点过程建模，其强度过程为 $\lambda(t,N_t)=(n-N_t)\mu_{x+t}$，其中 μ_{x+t} 是年龄为 $x+t$ 的死亡率.

无记忆的自激点过程的有限维分布 无记忆的自激点过程 N_t 在有限个时刻上的联合分布为：对于 $0\leqslant t_1<\cdots<t_m$，整数 $0\leqslant n_1<\cdots<n_m$，

$$P(N_{t_1}=n_1,\cdots,N_{t_m}=n_m)$$
$$=P_{\int_0^{t_1}\lambda(s,0)\mathrm{d}s}(n_1)P_{\int_{t_1}^{t_2}\lambda(s,n_1)\mathrm{d}s}(n_2-n_1)\cdots P_{\int_{t_{m-1}}^{t_m}\lambda(s,n_{m-1})\mathrm{d}s}(n_m-n_{m-1}),$$

其中

$$P_\mu(k)=\frac{\mu^k}{k!}\mathrm{e}^{-\mu}.$$

(2) 自激点过程的条件强度过程

定义 10.11

$$\hat{\lambda}(t,N_t)=E(\lambda_t\mid N_t) \tag{10.20}$$

称为自激点过程的**条件强度过程**.

于是，对于自激点过程有

$$P(N_{t+h}-N_t=1\mid N_t)=\hat{\lambda}(t,N_t)h+o(h),\quad P(N_{t+h}-N_t\geqslant 2\mid N_t)=o(h). \tag{10.21}$$

(3) 自激点过程的绝对概率

仿照初等概率论中 Poisson 过程的绝对概率的推导，可以得到

$$\begin{cases}P(N_t=0\mid N_s=0)=\mathrm{e}^{-\int_s^t\lambda(u,0)\mathrm{d}u},\\ P(N_t-N_s=0\mid N_s,\tau_1,\cdots,\tau_{N_s})=\mathrm{e}^{-\int_s^t\lambda(u,N_s;\,\tau_1,\cdots,\tau_{N_s})\mathrm{d}u}\quad (N_s\geqslant 1).\end{cases} \tag{10.22}$$

令

$$p_n(t)=P(N_t=n). \tag{10.23}$$

同样仿照初等概率论中 Poisson 过程的绝对概率的推导，我们得到下面的结论.

命题 10.9 自激点过程的绝对概率满足如下的(常微分方程组)

$$p_0'(t)=-\hat{\lambda}(t,0)p_0(t),$$
$$p_n'(t)=-\hat{\lambda}(t,n)p_n(t)+\hat{\lambda}(t,n-1)p_{n-1}(t)\quad (n\geqslant 1).$$

这是一个非时齐的单侧纯生过程. 可以用归纳法求得其解为

$$p_0(t)=\mathrm{e}^{-\int_0^t\hat{\lambda}(u,0)\mathrm{d}u}=\mathrm{e}^{-\int_s^t\lambda(u,0)\mathrm{d}u},$$
$$p_n(t)=\int_0^t\hat{\lambda}(s,n-1)p_{n-1}(s)\mathrm{e}^{-\int_s^t\hat{\lambda}(u,n)\mathrm{d}u}\mathrm{d}s\quad (n\geqslant 1).$$

(4) 自激点过程的事件到达时刻的联合分布与样本分布

现在我们求自激点过程 N_t 的前 n 个事件到达时刻$(\tau_1,\tau_2,\cdots,\tau_n)$的联合分布. 注意

对于$0<s_1<\cdots<s_n<t$ 有

$$\begin{aligned}\lambda(t,n;\ s_1,\cdots,s_n)h+o(h) &= P(N_{t+h}-N_t=1 \mid N_t=n,\tau_1=s_1,\cdots,\tau_n=s_n)\\ &= \frac{P(t<\tau_{n+1}\leqslant t+h, N_t=n \mid \tau_1=s_1,\cdots,\tau_n=s_n)}{P(N_t=n \mid \tau_1=s_1,\cdots,\tau_n=s_n)}\\ &= \frac{P(t<\tau_{n+1}\leqslant t+h \mid \tau_1=s_1,\cdots,\tau_n=s_n)}{P(\tau_{n+1}>t \mid \tau_1=s_1,\cdots,\tau_n=s_n)}.\end{aligned}$$

由此导出下述结论.

命题 10.10 $\lambda(t,n;\ s_1,\cdots,s_n)$是在 $\tau_1=s_1,\cdots,\tau_n=s_n$ 的条件下,τ_{n+1}的**条件故障率**. 因此,在 $\tau_1=s_1,\cdots,\tau_n=s_n$ 的条件下,τ_{n+1}的条件密度为

$$p_{\tau_{n+1}\mid\tau_1,\cdots,\tau_n}(s_{n+1} \mid s_1,\cdots,s_n) = \mathrm{e}^{-\int_{s_n}^{s_{n+1}}\lambda(u,n;\ s_1,\cdots,s_n)\mathrm{d}u}. \tag{10.24}$$

从而$(\tau_1,\tau_2,\cdots,\tau_n)$的联合分布为

$$p_{\tau_1,\cdots,\tau_n}(s_1,\cdots,s_n) = \prod_{k=1}^{n}\mathrm{e}^{-\int_{s_{k-1}}^{s_k}\lambda(u,k-1;\ s_1,\cdots,s_{k-1})\mathrm{d}u}, \tag{10.25}$$

其中 $s_0=0,\lambda(t,0;\ s_0)=\lambda(t,0)$.

定义 10.12 自激点过程在时刻 t 时的计数 N_t 与事件发生时刻的联合分布 $P(N_t=n,\tau_1\leqslant s_1,\cdots,\tau_n\leqslant s_n)$作为$(s_1,s_2,\cdots,s_n)$的函数的密度,称为**自激点过程的样本分布**.

我们可以进一步由(10.25)式,求出自激点过程的样本分布

$$\begin{aligned}p_{N_t,\tau_1,\cdots,\tau_{N_t}}(n,s_1,\cdots,s_n) &= P(N_t=n \mid \tau_1=s_1,\cdots,\tau_n=s_n)p_{\tau_1,\cdots,\tau_n}(s_1,\cdots,s_n)\\ &= P(N_t-N_{s_n}=0 \mid \tau_1=s_1,\cdots,\tau_n=s_n)\prod_{k=1}^{n}\mathrm{e}^{-\int_{s_{k-1}}^{s_k}\lambda(u,k-1;\ s_1,\cdots,s_{k-1})\mathrm{d}u}\\ &= \mathrm{e}^{-\int_{s_n}^{t}\lambda(u,n;\ s_1,\cdots,s_n)\mathrm{d}u}\prod_{k=1}^{n}\mathrm{e}^{-\int_{s_{k-1}}^{s_k}\lambda(u,k-1;\ s_1,\cdots,s_{k-1})\mathrm{d}u}.\end{aligned} \tag{10.26}$$

例 10.9 伽玛光子检测器用以检测按强度函数为 $v(t)$的非时齐的 Poisson 过程到达的伽玛光子. 然而,这种装置在检测到一个光子后有一个随机的失效时间. 假定检测到第 n 个光子后的失效时间为 ξ_n,$\{\xi_n\}$独立同分布,其共同分布函数是 $F_\xi(t)$,并且与此非时齐的 Poisson 过程独立. 如果将知道时刻 t 为止检测到的光子数记为 N_t,再记检测到光子的各个时刻分别依次为 $\tau_1,\tau_2,\cdots,\tau_n,\cdots$.

假定开始时检测装置是有效的. 我们说明 N_t 是自激点过程. 事实上,此时有 $\lambda(t,0)=v(t)$. 又因为在时刻 τ_n 检测到光子后,在其后的时刻 t 检测装置没有失效的概率是 $F_\xi(t-\tau_n-0)$(此处"-0"表示左极限). 于是对于 $N_t=n\geqslant1$ 有

$$\lambda(t,n;\ s_1,\cdots,s_n) = v(t)F_\xi(t-\tau_n-0) \quad (\text{即 } \lambda(t,N_t;\ \tau_1,\cdots,\tau_{N_t}) = v(t)F_\xi(t-\tau_{N_t}-0)).$$

特别地,当 ξ_n 为常数的时候,各个检测到光子的时刻$(\tau_1,\tau_2,\cdots,\tau_n)$的联合分布可以容易地由(10.26)式计算算得.

由(10.26)式直接得到下面的命题.

命题 10.11 自激点过程 N_t 是 m-记忆的，当且仅当$\{\tau_n\}$是 m 阶 Markov 链，即 $\boldsymbol{\xi}_n \stackrel{\text{def}}{=} \begin{pmatrix} \tau_{n+1} \\ \vdots \\ \tau_{n+m} \end{pmatrix}$是 Markov 链.

定义 10.13 1-记忆自激点过程 N_t 称为**时齐的**，如果存在一个函数 $h(x,s)$使

$$\lambda(t,N_t;\ \tau_{N_t}) = h(N_t, t-\tau_{N_t}).$$

于是可以推出下面的命题.

命题 10.12 假设 1-记忆自激点过程 N_t 的事件发生的时刻依次为$\{\tau_n\}$. 那么 N_t 是时齐的充要条件是$\{\tau_n\}$是独立随机变量序列的部分和.

10.5.2 带有随机调制的 Poisson 过程(二重 Poisson 过程)

定义 10.14 设$\{\xi_t\}$为随机过程(例如 Poisson 过程)，一般地称为**被调制的信息过程**. 而当 ξ_t 的轨道给定时，计数过程 N_t 是一个强度函数为 $\lambda(\xi_t)$的非时齐的 Poisson 过程，那么 N_t 称为**二重 Poisson 过程**，或**带随机调制**$\{\xi_t\}$的 Poisson 过程，$\lambda(\xi_t)$称为**二重 Poisson 过程的强度过程**.

注 强度过程也可以显含时间参数 t，即为 $\lambda(t,\xi_t)$的情形.

在实际问题中，常常是通过接收到的点过程 N_t 的一段样本轨道，来估计被调制的信息过程. 例如，信号为函数 $S(t)$，它在被均值为 0 的背景 Gauss 过程噪声的干扰下，成为代表信息的 Gauss 过程 ξ_t，而由此生成的二重 Poisson 过程 N_t 则是光子流. 最简单的情形是强度为 $\lambda(t,\xi_t)=\xi g(t)$(随机变量乘以常函数).

对于带随机调制$\{\xi_t\}$的 Poisson 过程 N_t，假定它依次的跳跃时刻为$\{\tau_n\}$，则 τ_n 们也都是$\{\xi_t\}$可知的，也就是说，它们都是$\{\xi_t\}$停时.

这时有

$$p_n(t) \stackrel{\text{def}}{=} P(N_t = n) = E\left[\mathrm{e}^{-\int_0^t \lambda(\xi_s)\mathrm{d}s} \frac{\left[\int_0^t \lambda(\xi_s)\mathrm{d}s\right]^n}{n!} \right].$$

下面的结论说明二重 Poisson 过程是自激 Poisson 过程的特殊情形.

命题 10.13 假定二重 Poisson 过程 N_t 的强度 $\lambda(\xi_t)$的数学期望有限. 记 N_t 的事件列分别依次为 $0<\tau_1<\cdots<\tau_n<\cdots$. 令

$$\lambda_t = E(\lambda(\xi_t) \mid N_u:\ u \leqslant t) = \begin{cases} E(\lambda(\xi_t) \mid N_t), & N_t = 0, \\ E(\lambda(\xi_t) \mid N_t, \tau_1, \cdots, \tau_{N_t}), & N_t \geqslant 1. \end{cases}$$

那么，N_t 是强度随机过程为 λ_t 的自激点过程.

证明 当 $h \to 0$ 时，有

$$E[P(N_{t+h} - N_t = 1 \mid N_s:\ s \leqslant t)] = E[P(N_{t+h} - N_t = 1 \mid N_s:\ s \leqslant t, \xi_t) \mid N_s:\ s \leqslant t]$$

$$= E[E(\lambda(\xi_t)h \mid N_s: s \leqslant t) \mid N_s: s \leqslant t] + o(h)$$
$$= \lambda_t h + o(h).$$

同样证明

$$P(N_{t+h} - N_t \geqslant 2 \mid N_s: s \leqslant t) = o(h).$$

于是二重 Poisson 过程的一些统计量的计算就直接化归自激点过程的计算. 而且相应的定理都可以应用.

注　一般自激点过程的 $\hat{\lambda}(t,n)$ 不易计算, 所以 $p_n(t)$ 也不易计算. 下面的定理可用于近似算法, 它可以利用特征函数方法证明.

命题 10.14(二重 Poisson 过程绝对概率的近似计算)　对二重 Poisson 过程 N_t, 记

$$\Lambda_t = \int_0^t \lambda(\xi_s)\mathrm{d}s, \quad N_t^* = \frac{N_t - EN_t}{\sqrt{\mathrm{var}(N_t)}}, \quad \Lambda_t^* = \frac{\Lambda_t - E\Lambda_t}{\sqrt{\mathrm{var}(\Lambda_t)}}.$$

假定 $t \to +\infty$ 时有 $E\Lambda_t \to +\infty$. 那么

(1) 在 $\dfrac{\mathrm{var}(\Lambda_t)}{(E\Lambda_t)^2} \to c > 0$ 时, 只要 Λ_t^* 依分布收敛于随机变量 η, N_t^* 就依分布收敛到 $\dfrac{\xi}{\sqrt{1+c}} + \dfrac{\eta}{\sqrt{1+\dfrac{1}{c}}}$, 其中 ξ 与 η 独立, 且 $\xi \sim N(0,1)$.

(2) 在 $\dfrac{\mathrm{var}(\Lambda_t)}{(E\Lambda_t)^2} \to 0$ 时, N_t^* 的分布收敛到 $N(0,1)$.

(3) 在 $\dfrac{\mathrm{var}(\Lambda_t)}{(E\Lambda_t)^2} \to +\infty$ 时, 只要 Λ_t^* 依分布收敛于随机变量 η, N_t^* 也依分布收敛于 η.

注　强度过程还可以显含时间 t, 即是 $\lambda(t,\xi_t)$ 的形式.

定义 10.15　设 ξ_t 是一个多维随机过程, 称为被调制的信息过程. 而当 ξ_t 的轨道给定时, 计数过程 $\mathbf{N}_t = (N_t^{(1)}, N_t^{(2)}, \cdots, N_t^{(d)})$ 的各个分量彼此独立, 且 $N_t^{(k)}$ 是一个强度函数为 $\lambda^{(k)}(\xi_t)$ 的非时齐的 Poisson 过程, 那么 $\mathbf{N}_t$ 称为**多道二重 Poisson 过程**. $\boldsymbol{\lambda}(\xi_t) = (\lambda^{(1)}(\xi_t), \cdots, \lambda^{(d)}(\xi_t))$ 称为**多道二重 Poisson 过程的强度过程**.

在应用中, 如核医疗中多探针的检测系统、伽玛射线照相机、核磁共振等, 常用多道二重 Poisson 过程建模. 更多地出现的是 ξ_t 是 Markov 过程, 或 Markov 链的情形. 用 $\mathbf{N}_t$ 的观测样本去估计信息过程 ξ_t 就是滤波.

10.5.3　滤波 Poisson 过程

定义 10.16　设 N_t 是强度为 $\lambda(t)$ 的非时齐的 Poisson 过程, $\{\eta_n\}$ 是与之独立的独立同分布随机变量序列. 又设 N_t 的第 n 个事件发生时刻 τ_n 的所负的代价不是由 η_n 直接给出, 而是通过某个有界的响应函数 $h(t,s,x)(s \leqslant t)$ 给出的 $h(t,\tau_n,\eta_n)$. 于是累计能量为

$$\xi_t = \int_0^t h(t, s, \widetilde{\eta}_s) \mathrm{d}N_s,$$

$$\widetilde{\eta}_s = \begin{cases} \eta_n, & s = \tau_n, \\ 0, & \text{其他}. \end{cases}$$

称 ξ_t 为**滤波 Poisson 过程**. 特别地，当 $h(t,\tau_n,\eta_n)=\eta_n$ 时，滤波 Poisson 过程就是非时齐的复合 Poisson 过程(所以滤波 Poisson 过程的含义就是“扭曲了的非时齐的复合 Poisson 过程”).

10.6 节介绍特征泛函，这是一个有力的数学工具. 也适合于 Gauss 过程等其他随机过程.

10.6　特征泛函

随机过程的特征泛函相应于随机变量或随机向量的特征函数或联合特征函数. 特征泛函既等价于整个过程的分布 (有限维分布函数族)，又能快速的得到一切矩.

10.6.1　Poisson 过程的特征泛函

定义 10.17　对于 Poisson 过程 N_t 及定义在 $[0,T]$ 上的函数 $f(t)$，我们将 $\int_0^T f(s)\mathrm{d}N_s$ 的特征函数在 1 处的值记为 $\Phi_N(f)$，即

$$\Phi_N(f) = E\mathrm{e}^{\mathrm{i}\int_0^T f(s)\mathrm{d}N_s}.$$

这里的 f 是参变函数，对于给定一个函数 f，就有一个数 $\Phi_N(f)$ 与之对应，这种从函数 f 到 $\Phi(f)$ 的映射称为**泛函**，又因为此泛函是通过 Poisson 过程的积分生成的，所以称为 **Poisson 过程的特征泛函**.

例 10.10　当 $f(s)\equiv\theta I_{[0,t]}(s)$ 时，Poisson 过程的特征泛函就简化为 Poisson 过程在时刻 t 的特征函数 $E\mathrm{e}^{\mathrm{i}\theta N_t}$. 而当 $f(t)\equiv\theta_1 I_{[0,t_1]}(t)+\theta_2 I_{(0,t_2]}(t)$ 时，特征泛函就简化为 Poisson 过程在时刻 t_1 与时刻 t_2 的联合特征函数 $E\mathrm{e}^{\mathrm{i}(\theta_1 N_{t_1}+\theta_2 N_{t_2})}$.

特征泛函是有限个时刻的联合特征函数在一个区间上的所有时刻情形的自然推广. 为了将两个时刻的联合特征函数推广到反映时间区间 $[0,T]$ 中一切时刻的特征泛函，我们需要改写两个时刻的联合特征函数成为可以推广的形式如下.

假设 $t_1<t_2\leqslant T$. 利用 Poisson 过程的独立增量性与时齐性，对于 $f(t)=\theta_1 I_{[0,t_1]}(t)+\theta_2 I_{(t_1,t_2]}(t)$，有

$$\begin{aligned}\Phi(f) &= \Phi(\theta_1 I_{[0,t_1]}+\theta_2 I_{(t_1,t_2]}) = E\mathrm{e}^{\mathrm{i}(\theta_1 N_{t_1}+\theta_2(N_{t_2}-N_{t_1}))} = E\mathrm{e}^{\mathrm{i}\theta_1 N_{t_1}} E\mathrm{e}^{\mathrm{i}\theta_2 N_{t_2-t_1}} \\ &= \mathrm{e}^{\lambda t_1(\mathrm{e}^{\mathrm{i}\theta_1}-1)}\mathrm{e}^{\lambda(t_2-t_1)(\mathrm{e}^{\mathrm{i}\theta_2}-1)} = \mathrm{e}^{\int_0^T \lambda(\mathrm{e}^{\mathrm{i}[\theta_1 I_{[0,t_1]}(t)+\theta_2 I_{(t_1,t_2]}(t)]}-1)\mathrm{d}t} = \mathrm{e}^{\int_0^T \lambda(\mathrm{e}^{\mathrm{i}f(t)}-1)\mathrm{d}t}.\end{aligned}$$

这是推广的起点. 再复杂一些，对于 $0=t_0<t_1<\cdots<t_n\leqslant T$，及 $f(t)=\sum_{k=0}^{n-1} c_k I_{(t_k,t_{k+1}]}(t)$，

我们类似地可以得到

$$Ee^{i\int_0^T f(t)\mathrm{d}N_t} = \Phi_N(f) = e^{\int_0^T \lambda(e^{if(t)}-1)\mathrm{d}t}.$$

这个公式对于任意连续函数,甚至更为一般的函数,可以利用函数近似的方法证明它仍然正确.于是我们得到如下的定理.

定理 10.15 Poisson 过程的特征泛函的表达公式为

$$Ee^{i\int_0^T f(t)\mathrm{d}N_t} = \Phi_N(f) = e^{\int_0^T \lambda(e^{if(t)}-1)\mathrm{d}t}. \tag{10.27}$$

10.6.2 非时齐的 Poisson 过程的特征泛函

定义 10.17′ 非时齐的 Poisson 过程的特征泛函仍定义为

$$\Phi_N(f) = Ee^{i\int_0^T f(t)\mathrm{d}N_t}.$$

类似地得到下面的定理.

定理 10.15′ 非时齐的 Poisson 过程的特征泛函的表达公式为

$$Ee^{i\int_0^T f(t)\mathrm{d}N_t} = \Phi_N(f) = e^{\int_0^T \lambda(t)(e^{if(t)}-1)\mathrm{d}t}. \tag{10.27′}$$

推论 10.16 鞅 $\widetilde{N}_t = N_t - \int_0^T \lambda(s)\mathrm{d}s$ 的特征泛函为

$$\Phi_{\widetilde{N}}(f) \stackrel{\text{def}}{=} Ee^{i\int_0^T f(s)\mathrm{d}\widetilde{N}_s} = e^{\int_0^T [e^{if(s)}-1-if(s)]\lambda(s)\mathrm{d}s}. \tag{10.28}$$

推论 10.17

(1) $E(\widetilde{N}_s\widetilde{N}_t) = \mathrm{cov}(N_s, N_t) = \int_0^{s\wedge t}\lambda(u)\mathrm{d}u.$

(2) $E(\widetilde{N}_{t_1}\widetilde{N}_{t_2}\widetilde{N}_{t_3}) = \int_0^{t_1\wedge t_2\wedge t_3}\lambda(u)\mathrm{d}u.$

(3) $E(\widetilde{N}_{t_1}\widetilde{N}_{t_2}\widetilde{N}_{t_3}\widetilde{N}_{t_4}) = \int_0^{t_1\wedge t_2\wedge t_3\wedge t_4}\lambda(u)\mathrm{d}u + \mathrm{cov}(N_{t_1}, N_{t_2})\mathrm{cov}(N_{t_3}, N_{t_4})$
$+ \mathrm{cov}(N_{t_1}, N_{t_3})\mathrm{cov}(N_{t_2}, N_{t_4}) + \mathrm{cov}(N_{t_1}, N_{t_4})\mathrm{cov}(N_{t_2}, N_{t_3}).$

证明 取 $f(t) = \sum_{k=1}^4 \theta_k I_{[0,t_k]}(t)$. 将 $\Phi_{\widetilde{N}}(f)$ 记为 $\varphi(\theta_1, \theta_2, \cdots, \theta_4)$. 求 $\frac{\partial^4\varphi}{\partial\theta_1\cdots\partial\theta_4}(0,0,0,0)$,就得到结论(3). 其他的结论可类似证得.

注 对于一般未必可微的递增函数 Λ_t,我们可以推广地定义以 Λ_t 为补偿函数的非时齐的 Poisson 过程 N_t 为:非时齐的独立增量过程,且对于任意 $s<t$,有 $N_t - N_s \sim \mathrm{Poisson}_{\Lambda_t-\Lambda_s}$. 此时相应的结论都可以推广.

10.6.3 非时齐的复合 Poisson 过程的特征泛函

命题 10.18 定义非时齐的复合 Poisson 过程 ξ_t 在区间$[0,T]$上的特征泛函为

$$\Phi_\xi(f) \stackrel{\text{def}}{=} Ee^{i\int_0^T f(t)\mathrm{d}\xi_t}.$$

其中积分的含义如(10.3)式,即 $\xi_t=\int_0^t \widetilde{\eta}_s \mathrm{d}N_s$,而 $\widetilde{\eta}_s=\begin{cases}\eta_n, & s=\tau_n,\\ 0, & \text{其他}.\end{cases}$

那么它的明显表达式为

$$\Phi_\xi(f)=\mathrm{e}^{\int_0^T[\varphi(f(u)-1)]\lambda(u)\mathrm{d}u}, \tag{10.29}$$

其中 $\varphi(\theta)$是 η_1 的特征函数:$\varphi(\theta)=E\mathrm{e}^{\mathrm{i}\theta\eta_1}$.

证明　与(10.27)式类似地有

$$E[\mathrm{e}^{\mathrm{i}\sum_{k=1}^{N_T} f(\tau_k)X_k} \mid N_T=n]=E\mathrm{e}^{\mathrm{i}\sum_{k=1}^{n} f(\tau_k)\eta_k}=E\left[\prod_{k=1}^{n}\varphi(f(\eta_k))\right]$$
$$=\frac{1}{\left(\int_0^T \lambda(u)\mathrm{d}u\right)^n}\left[\int_0^T \varphi(f(u))\lambda(u)\mathrm{d}u\right]^n.$$

再用全期望公式得到

$$\begin{aligned}\Phi_Y(f)&=E\mathrm{e}^{\mathrm{i}I_{\{k\geqslant 1\}}\sum_{k=1}^{N_T} f(\tau_k)\eta_k}\\&=P(N_T=0)+\sum_{n=1}^{+\infty}E(\mathrm{e}^{\mathrm{i}\sum_{k=1}^{n} f(\tau_k)\eta_k} \mid N_T=n)P(N_T=n)\\&=\mathrm{e}^{-\int_0^T\lambda(u)\mathrm{d}u}\left[1+\sum_{n=1}^{+\infty}\frac{1}{\left(\int_0^T\lambda(u)\mathrm{d}u\right)^n}\left[\int_0^T\varphi(f(u))\lambda(u)\mathrm{d}u\right]^n\frac{\left(\int_0^T\lambda(u)\mathrm{d}u\right)^n}{n!}\right]\\&=\mathrm{e}^{-\int_0^T\lambda(u)\mathrm{d}u}\mathrm{e}^{\int_0^T\varphi(f(u))\lambda(u)\mathrm{d}u}=\mathrm{e}^{\int_0^T[\varphi(f(u))-1]\lambda(u)\mathrm{d}u}.\end{aligned}$$

因为具有相同的特征泛函的两个随机过程的统计性质是一样的,所以命题 10.18 可以用于由特征泛函的形式来确定一个随机过程是否是非时齐的复合 Poisson 过程.

例 10.11(非时齐的广义 Poisson 过程)　如果赋值随机变量具有离散分布 $\eta_n\sim\begin{pmatrix}1 & \cdots & k & \cdots\\ p_1 & \cdots & p_k & \cdots\end{pmatrix}$,则非时齐的复合 Poisson 过程称为**非时齐的广义 Poisson 过程**. 此时的特征泛函为

$$\Phi_Y(f)=\mathrm{e}^{\int_0^T\left[\sum_{k=1}^{+\infty}p_k\mathrm{e}^{\mathrm{i}f(u)k}-1\right]\lambda(u)\mathrm{d}u}. \tag{10.30}$$

这个公式可以用命题 10.18 类似的方法证明.

定义 10.18　滤波 Poisson 过程 ξ_t 在区间$[0,T]$上的特征泛函定义为:对于任意连续的增函数 $F(t)$,令

$$\bar{\Phi}_\xi(F)\overset{\mathrm{def}}{=\!=}E\mathrm{e}^{\mathrm{i}\int_0^T\xi_t\mathrm{d}F(t)}, \tag{10.31}$$

(用$\int\xi_t\mathrm{d}F(t)$ 形式的积分定义特征泛函要比用$\int f(t)\mathrm{d}\xi_t$ 形式的积分更为灵活,使用面更

广$\Big)$. 类似地还可以证明下面的命题.

命题 10.19 滤波 Poisson 过程 ξ_t 在区间$[0,T]$上的特征泛函有如下的表达式：

$$\bar{\Phi}_{\xi}(F) = e^{\int_0^T \left[E e^{i\int_u^T h(v,u,\eta_1)\mathrm{d}F(v)} - 1 \right] \lambda(u)\mathrm{d}u}. \tag{10.32}$$

由此得到

$$E\xi_t = \int_0^t E[h(t,u,\eta_1)]\lambda(u)\mathrm{d}u, \quad \mathrm{cov}(\xi_s,\xi_t) = \int_0^{s\wedge t} E[h(s,u,\eta_1)h(t,u,\eta_1)]\lambda(u)\mathrm{d}u.$$

用特征函数方法还不难证明下述类似于中心极限定理的近似定理.

命题 10.20(时齐的滤波 Poisson 过程的泛函型的 Gauss 过程近似定理) 假定生成滤波 Poisson 过程的 Poisson 过程是时齐的，其强度为 λ. 那么，当 $\lambda\to+\infty$时，在$(0,T]$上 $\dfrac{\xi_t - E\xi_t}{\sqrt{\mathrm{var}(\xi_t)}}$依分布收敛到 ζ_t，其中 ζ_t 是期望函数为 0 的 Gauss 过程，其协方差函数为 $\mathrm{cov}(\zeta_s,\zeta_t)=\rho_{\xi_s,\xi_t}$($\xi_s$ 与 ξ_t 的相关系数).

使用这个定理，可以将高强度的时齐的滤波 Poisson 过程的近似计算大大地简化.

10.7 随机地模拟 Poisson 过程和非时齐的 Poisson 过程的方法

10.7.1 模拟 Poisson 过程的方法

我们首先模拟一个均值为 λt 的 Poisson 随机变量 $N(t)$. 如果 $N(t)=n$，再生成 n 个随机数：$U_1,U_2,\cdots,U_n$，将 $U_1,U_2,\cdots,U_n$ 排序，得到 $U_{(1)},U_{(2)},\cdots,U_{(n)}$，于是 $tU_{(1)},tU_{(2)},\cdots,tU_{(n)}$ 就表示 n 个相继到达的事件的时间.

10.7.2 模拟非时齐的 Poisson 过程

设 N_t 是强度函数 $\lambda(t)(0\leqslant t<+\infty)$的非时齐的 Poisson 过程，要模拟这个过程在时间 T 以前所发生的事件，文献[Ro]中 Ross 提出了三种模拟方法. 本书只介绍最简便的**采样于 Poisson 过程的削薄算法.**

削薄算法的做法是：寻找 λ(越小越好)使

$$\lambda(t)\leqslant\lambda \quad 任意\ t\leqslant T$$

强度函数 $\lambda(t)$的非时齐的 Poisson 过程可以由速率 λ 的 Poisson 过程的事件的时间，经过一个随机选取得到. 即对速率 λ 的 Poisson 过程在时间 t 发生的一个事件以概率 $\lambda(t)/\lambda$ 被计数，那么被计数的事件是一个强度函数为 $\lambda(t)(0\leqslant t<+\infty)$的非时齐的 Poisson 过程.

削薄算法可以改进为：将区间分裂割为子区间，而后在每个子区间上用这个程序. 就

是说，确定合适的值 k，$0<t_1<t_2<\cdots<t_k<T$，$\lambda_1,\cdots,\lambda_{k+1}$ 使

$$\lambda(s)\leqslant\lambda_i\quad 当\ t_{i-1}\leqslant s<t_i,\quad i=1,\cdots,k+1\quad (其中\ t_0=0, t_{k+1}=T).\qquad(10.33)$$

现在，在区间 $[t_{i-1},t_i)$ 通过生成速率为 λ_i 的指数随机变量，并且以概率 $\lambda(s)/\lambda_i$ 接受在时间 s，$s\in[t_{i-1},t_i)$ 发生的事件. 更精细的算法可以参见文献[Ro].

习题 10

1. 假设公共汽车站在时刻 T 发车，而顾客则以参数为 λ 的指数流来到车站.

(1) 证明在发车前的所有顾客的等待时间的和的平均值是 $\frac{1}{2}\lambda T^2$.

(2) 如果在时刻 T 前的某个固定时刻 s 增加一班汽车. 证明如果选取 $s=\frac{T}{2}$，那么在时刻 T 前到达车站的所有顾客的平均总等待时间为最小.

(3) 如果在问题(1)中在时刻 T 前的某个停时 τ 的时刻增加一班汽车. 证明在时刻 T 前到达车站的所有顾客的平均总等待时间为 $\frac{1}{2}\lambda T^2-E[(T-\tau)N_\tau]$.

(4) 证明(3)中有 $E[(T-\tau)N_\tau]=E\int_0^\tau[\lambda(T-t)-N_t]\mathrm{d}t$.

(5) 证明(3)中如取 $\tau=\inf\{t: N_t\geqslant\lambda(T-t)\}$，则所有顾客的平均总等待时间最小.

(6) 证明(3)中在 $[0,t]$ 中的累计到达顾客数 N_t 是强度为 $\lambda(t)$ 的非时齐 Poisson 过程，并且如取 $\tau=\inf\{t: N_t\geqslant\lambda(t)(T-t)\}$，则所有顾客的平均总等待时间最小.

2. 证明非时齐的复合 Poisson 过程是非时齐的独立增量过程. 再求它的数学期望函数与协方差函数.

3. 如果赋值随机变量具有分布 $\eta_n\sim\begin{pmatrix}0 & 1\\ 1-p & p\end{pmatrix}$. 这时的非时齐的复合 Poisson 过程是什么? 解释其概率含义.

4. 设 $f(t)$ 是数值函数，N_t 是非时齐的 Poisson 过程. ξ_t 满足随机微分方程

$$\mathrm{d}\xi_t=\xi_t f(t)\mathrm{d}N_t,\quad \xi_0=1.$$

(1) 证明 $E\xi_t=\mathrm{e}^{\int_0^t f(s)\lambda(s)\mathrm{d}s}$.

(2) 求 $\mathrm{d}\xi_t^2$.

(3) 求 $\mathrm{var}(\xi_t)$，$\mathrm{var}(\xi_t^2)$.

5. 证明自激点过程 N_t 的数学期望和方差分别为

$$EN_t=\int_0^t(E\lambda_u)\mathrm{d}u,$$

$$\mathrm{var}(N_t)=\int_0^t (E\lambda_u)\mathrm{d}u+2\int_0^t E(N_u\lambda_u)\mathrm{d}u-\left[\int_0^t (E\lambda_u)\mathrm{d}u\right]^2,$$

其中随机过程 λ_t 是 N_t 的随机强度过程.

6. 假设 0-记忆自激点过程 N_t 满足：$\lambda(N_t)=\lambda(0)[N_t+1]$. 证明

$$p_n(t)=p_0(t)[1-p_0(t)]^n.$$

7. 对于二重 Poisson 过程 N_t，记 $\Lambda_t=\int_0^t \lambda(s,\xi_s)\mathrm{d}s$. 证明

$$\mathrm{cov}(N_s,N_t)=\mathrm{cov}(\Lambda_s,\Lambda_t)+E\Lambda_{s\wedge t}.$$

8. 设二重 Poisson 过程 N_t 的 $\lambda(t,\xi_t)=g(t)\xi$，其中 $g(t)$ 非随机，ξ 有分布密度 $p_\xi(t)I_{[0,\infty)}(t)$. 证明

$$P(N_t=n)=\frac{\left[\int_0^t g(s)\mathrm{d}s\right]^{nt}}{n!}\int_0^{+\infty} t^n \mathrm{e}^{-t\int_0^t g(s)\mathrm{d}s} p_\xi(t)\mathrm{d}t.$$

$$\hat{\lambda}(t,N_t)=E(\lambda_t\mid N_t)=E[E(\lambda(t,\xi_t)\mid N_s:s\leqslant t)\mid N_t]$$

$$=E[\lambda(t,\xi_t)\mid N_t]=\frac{g(t)\int_0^{+\infty} s^{N_t+1}\mathrm{e}^{-s\int_0^t g(u)\mathrm{d}u} p_\xi(s)\mathrm{d}s}{\int_0^{+\infty} s^{N_t}\mathrm{e}^{-s\int_0^t g(u)\mathrm{d}u} p_\xi(s)\mathrm{d}s}.$$

如果 $\xi\sim$Gamma 分布 $\Gamma(\alpha,\lambda)$，那么

$$P(N_t=n)=\frac{\Gamma(n+\alpha)}{n!\Gamma(\alpha)}\left[\frac{\lambda}{\lambda+\int_0^t g(s)\mathrm{d}s}\right]^\alpha\left[\frac{\int_0^t g(s)\mathrm{d}s}{\lambda+\int_0^t g(s)\mathrm{d}s}\right]^n.$$

N_t 称为**非时齐的 Polya 过程**.

名词索引

参 考 文 献

[A] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. John Wiley & Sons,1974.

[BW] R. M. Bhattacharya,W. C. Waimire. *Stochastic Processes with Applications*. John Wiley & Sons, 1990.

[C] K. L. Chung(钟开莱). *Elementary Probability Theory with Stochastic Processes*. Springer,1975. 中译本：概率论附随机过程. 人民教育出版社,1980.

[E] Evans, Lawrence C. *An Introduction to Stochastic Differential Equations. Version* 1. 2, Lecture Notes of Department of Mathematics UC Berkeley.

[GQ] 龚光鲁,钱敏平. 应用随机过程教程：与在算法和智能计算中的应用. 北京：清华大学出版社,2004.

[GS] I. I. Gihman,A. V. Skorohod. *Stochastic Differential Equations*. Springer,1972.

[H] 何声武. 随机过程论. 上海：华东师范大学出版社,1989.

[Hi] D. J. Higham. *An algorithmic introduction to numerical simulation of stochastic differential equations*. SIAM Review **43** (2001),525-546.

[L1] J. Lamperti. *Probability*. Benjamin,1968.

[L2] J. Lamperti. *A simple construction of certain diffusion processes*. J. Math. Kŷoto,1964,161-170.

[Li] M. van Lieshout. *Markov Point Processes and their Applications*. Imperial College Press, The Nietherland,2000. (distributed by World Scientific Publishing Co,Singapore).

[N] Salih N. Neftci. *An Introduction to the Mathematics of Financial Derivatives*. 2nd ed. Academic Press. 2000.

[Ø] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. 4th ed. Springer,1995.

[PW] E. Platen, W. Wagner. On a Taylor formula for a class of Ito process. Probability Math. Statistics 3,37-51,1982.

[QG] 钱敏平，龚光鲁. 应用随机过程. 北京：北京大学出版社,1998.

[QY] 钱敏平，叶俊. 随机数学. 第 2 版. 北京：高等教育出版社,2004.

[Ro] S. Ross. *Introduction to Stochastic Models*. 第 9 版. Academic Press,2006. (有邮电出版社的影印本与中译本).

[S] Z. Schuss. *Theory and Applications of Stochastic Differential Equations*. John Wiley & Sons,1980.

[Sn] D. L. Snyder. 随机点过程. (梁之舜,邓永录译). 北京：人民教育出版社,1981.

[Su1] H. Sussmann. *An interpretation of stochastic differential equations as ordinary differential equations which depend on the sample point*. Bulletin AMS **83** (1977),296-298.

[Su2] H. Sussmann. *On the gap between deterministic and stochastic ordinary differential equations*. Ann. Probability **6** (1978), 19-41.

[T] H. C. Tijms. Stochastic Models, An Algorithmic Approach. John Wiley & Sons, 1994.

[Wg] 王梓坤. 随机过程论. 北京：科学出版社，1965.

[YWL] 严士健，王隽骧，刘秀芳. 概率论基础. 北京：科学出版社，1982.